THE UNIVERSITY OF
WARWICK

University of Warwick institutional repository: http://go.warwick.ac.uk/wrap

A Thesis Submitted for the Degree of PhD at the University of Warwick

http://go.warwick.ac.uk/wrap/35249

This thesis is made available online and is protected by original copyright.

Please scroll down to view the document itself.

Please refer to the repository record for this item for information to help you to cite it. Our policy information is available from the repository home page.

THE UNIVERSITY OF WARWICK

AUTHOR: **Nastasiya Grinberg** DEGREE: **Ph.D.**

TITLE: **Some problems in stochastic analysis: Itô's formula for convex functions, interacting particle systems and Dyson's Brownian motion**

DATE OF DEPOSIT: ....

I agree that this thesis shall be available in accordance with the regulations governing the University of Warwick theses.

I agree that the summary of this thesis may be submitted for publication.

I **agree** that the thesis may be photocopied (single copies for study purposes only).

Theses with no restriction on photocopying will also be made available to the British Library for microfilming. The British Library may supply copies to individuals or libraries. subject to a statement from them that the copy is supplied for non-publishing purposes. All copies supplied by the British Library will carry the following statement:

> "Attention is drawn to the fact that the copyright of this thesis rests with its author. This copy of the thesis has been supplied on the condition that anyone who consults it is understood to recognise that its copyright rests with its author and that no quotation from the thesis and no information derived from it may be published without the author's written consent."

AUTHOR'S SIGNATURE: ....

USER'S DECLARATION

1. I undertake not to quote or make use of any information from this thesis without making acknowledgement to the author.
2. I further undertake to allow no-one else to use this thesis while it is in my care.

DATE SIGNATURE ADDRESS

....

....

....

....

....

Some problems in stochastic analysis: Itô's formula for convex functions, interacting particle systems and Dyson's Brownian motion

by

Nastasiya Grinberg

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

Department of Statistics

February 2011

THE UNIVERSITY OF
WARWICK

Contents

Acknowledgments

First of all I would like to thank my supervisors Prof. Wilfrid Kendall and Dr. Jon Warren, both of whom have been infinitely patient with me all this time, and whose invaluable help allowed me to finish this exciting but tough journey. I will especially hold dear the pre-submission e-mail frenzy that I have subjected Jon to, in what seems like, in vain.

I am also grateful to all of the friends I have made in the Statistics Department and especially to my housemates, Peter and Martin, all of whom supported and encouraged me when I most needed it.

Needless to say that I would not have been where I am now if not for my parents and my sister and their unconditional love and support. Doing most things in life would have been pointless if not for my family.

Lastly, I would like to acknowledge the financial support of the Statistics Department which made doing a PhD a reality.

Declarations

I hereby declare that this thesis is based on my own work, except where stated otherwise, and has not been submitted for a degree at any other university. Material appearing in Chapter 1 is based on [38] and has been submitted to *ESAIM: Probability and Statistics*.

Abstract

This thesis consists of two main parts: Chapter 1 is concerned with studying an extension of the Itô lemma to the convex functions. We prove that the local martingale part of the decomposition of a convex function f of a continuous semimartingale can be expressed in a similar way to the classical formula with the gradient of f replaced with its *subgradient*. The result itself is not new, however, our approach via Brownian perturbation is.

The second, and the largest, part of the thesis focusses on the study of a certain family of bivariate diffusions $Z^{(\theta,\mu)} = (X, R)$ in a wedge $W = \{(x, r) \in \mathbb{R} \times \mathbb{R}^+ : |x| \leq r\}$, parameterised by $\theta \in (0, \infty)$ and $\mu \geq 0$, with the property that X is distributed as a Brownian motion with drift μ and R is the so-called 3-dimensional Bessel process of drifting Brownian motion. By letting parameter θ tend to ∞ and 0 we can recover the two well-known couplings of the two processes coming from the Pitman's theorem and by considering radial part of the 3-dimensional BM (with drift $\mu \geq 0$) respectively. This family of continuous processes is obtained as a diffusion approximation in Chapter 3 of a certain family of two-dimensional Markov chains arising in representation theory and is characterised, for each $\theta \in (0, \infty)$ and $\mu \geq 0$, as a unique solution to a certain martingale problem in Chapter 4. Moreover, we show that the process $Z^{(\mu,\theta)}$ together with the marginal R-process provide an example of *intertwined* processes. Finally, in Chapter 5 we consider a family of certain Markov chains in the Gelfand-Cetlin cone of depth n. We show that for $n = 2$ the Markov chains of Chapter 3 can be recovered. We identify several intertwining relationships and make a step towards linking the diffusion limit of the chain to a certain Markov function of the GUE minor process of random

matrix theory, which consists of two interlaced Dyson's Brownian motions.

Chapter 1

Itô's formula for convex functions of continuous semimartingales: Brownian perturbation approach.

The first chapter of this Thesis concentrates on the study of convex functions in semimartingale theory; it is self-contained and covers a topic different from the rest of the Thesis.

We consider an extension of the celebrated Itô's formula for the continuous semimartingales to the class of convex functions. In particular we prove that the local martingale part of a convex function f of a d-dimensional semimartingale $X = M + A$ can be written in terms of an Itô stochastic integral $\int H(X)dM$, where $H(x)$ is some particular measurable choice of subgradient $\overline{\nabla} f(x)$ of f at x, and M is the martingale part of X. This result was first proved by Bouleau in [15]. Here we present a new treatment of the problem. We first prove the result for $\widetilde{X} = X + \epsilon B$, $\epsilon > 0$, where B is a standard Brownian motion, and then pass to the limit as $\epsilon \to 0$, using results of Barlow and Protter [4] and Carlen and Protter [19]. The former paper concerns convergence of semimartingale decompositions of semimartingales, while the latter studies a special case of converging convex functions of semimartingales. Material from this chapter appears in [38].

1.1 Classical Itô's formula and its extensions

Let X be a continuous $\mathbb{R}^d$-valued semimartingale with decomposition $X_t = X_0 + M_t + A_t$, where M is a local martingale and A is a process of finite variation, and let f be a twice continuously differentiable function. Then Itô's lemma states that $f(X)$ is also a continuous semimartingale and that, moreover,

$$f(X_t) = f(X_0) + \sum_i \int_0^t \frac{\partial f}{\partial x_i}(X_s)dX_s + \frac{1}{2}\sum_{i,j}\int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(X_s)d\langle X_s, X_s\rangle \ .$$

In particular, the martingale part of $f(X)$ is given by

$$N_t := \sum_i \int_0^t \frac{\partial f}{\partial x_i}(X_s)dM_s = \int_0^t \nabla f(X_s)dM_s \ ,$$

where $\nabla f(x)$ is the gradient of f at x (see section 1.3.1 for definition and more details).

One of the most well-known, and also one of the simplest, extensions of the Itô's formula is the *Tanaka's formula* (see, for example, [64, p. 169]), which states that for a standard Brownian motion B we have

$$|B_t| = |B_0| + \int_0^t \operatorname{sgn}(B_s)dB_s + L_t^0 \ ,$$

where $\operatorname{sgn}(x) = -1$ for $x \le 0$ and $\operatorname{sgn}(x) = 1$ for $x > 0$, and L_t is the so called *local time* of B at the origin, given by

$$L_t^0 = \lim_{\epsilon \downarrow 0} \frac{1}{2\epsilon} |\{s \le t : B_s \in (-\epsilon, +\epsilon)\}| \ .$$

Note that Tanaka's formula is also the easiest extension of Itô's result to convex functions. Meyer proves a more general result

Theorem 1.1. (Meyer *[57, p. 361]*) *Let X be an $\mathbb{R}^d$-valued semimartingale and let f be a convex function on $\mathbb{R}^d$. Then $f(X)$ is again a semimartingale.*

Moreover, for the case $d = 1$ he derives the change of variables formula

$$f(X_t) = f(X_0) + \int_0^t \nabla_- f(X_s) dX_s + \int_{\mathbb{R}} L_t^a \mu(da) ,$$

where $\nabla_- f$ is the left-hand side derivative of f, and $\mu = \frac{1}{2}\frac{\partial^2 f}{\partial x^2}$ is viewed as a measure. The increasing process L_t^a is the *local time process of* X given by

$$L_t^a(X) = \lim_{\epsilon \downarrow 0} \frac{1}{\epsilon} \int_0^t \mathbf{1}_{[a,a+\epsilon)}(X_s) d\langle X_s, X_s \rangle ,$$

and which can be viewed as the time X spends at a. We refer to the above formula as the *Meyer-Itô formula*. The so-called *Meyer-Tanaka formula* is its special case: for a continuous semimartingale X and a function $f(x) = |x|$ we have

$$|X_t| = |X_0| + \int_0^t \operatorname{sgn}(X_s) dX_s + L_t^0 .$$

When X is standard Brownian motion, we recover Tanaka's formula.

Now consider a general convex function $f : \mathbb{R}^d \to \mathbb{R}$, not necessarily everywhere differentiable. Every differentiable point $x \in \mathbb{R}^d$ has a unique tangential hyperplane, while at non-differentiable points there is a whole set of supporting hyperplanes. For a continuous semimartingale X with decomposition $X = M + A$ we prove that the (local) martingale part of $f(X)$ can be expressed in terms of a stochastic integral of a measurable selection of subgradient $\overline{\nabla} f(X)$ against M. For piecewise linear 1-dimensional convex functions this follows from the Meyer-Tanaka formula. For example, for $f(x) = |x|$ we have $\overline{\nabla} f(x) = \operatorname{sgn}(x)$. So at the origin, which is the only point where derivative is not defined, we can take the supporting line to be $y = -x$. Moreover, because as we know Brownian motion spends zero time in Lebesgue-null sets, we can in fact choose $\overline{\nabla} f(0)$ to be any number in the interval $[-1, +1]$ (corresponding to the possible slopes of supporting lines at 0).

For a continuous semimartingale X with decomposition $X = M + A$ we define

$$\| X \|_{\mathscr{H}^p} = \| \langle M, M \rangle_\infty^{1/2} + \int_0^\infty |dA_s| \|_{L^p}$$

for $1 \le p < \infty$. The $\mathscr{H}^p$-space consists of all semimartingales X such that $\| X \|_{\mathscr{H}^p} < \infty$. The main result of this chapter is the following

Theorem 1.2. *Let $f : \mathbb{R}^d \to \mathbb{R}$ be convex and let X be a continuous $\mathbb{R}^d$-valued semimartingale defined on filtered probability space $(\Omega, \mathscr{F}, \{\mathscr{F}\}_{t \ge 0}, \mathbb{P})$ with Meyer decomposition $X_t = X_0 + M_t + A_t$. Then $f(X_t)$ is again a continuous semimartingale; in particular, its local martingale part is given by*

$$\int_0^t \overline{\nabla} f(X_s) dM_s \ , \quad \textit{locally in } \mathscr{H}^1 \ ,$$

where $\overline{\nabla} f(x)$ is some choice of subgradient of f at x, such that $\overline{\nabla} f(X_t)$ is $\mathscr{F}_t$-adapted.

The above theorem was first stated and proved by Bouleau ([15, Lemma 2]). In the follow-up paper [16] he proves the conjecture stated in [15] that in fact *any measurable choice* of $\overline{\nabla} f(x)$ can be used. In this chapter we prove the first of the two results using an approach different to that in [15].

There are many other papers on extending the Itô's formula by considering different classes of functions f or stochastic processes, or both. In [72], for example, Russo and Vallois derive Itô's formula for $\mathscr{C}^1(\mathbb{R}^d)$-functions of continuous semimartingales whose time-reversals are also continuous semimartingales. They also extend the formula to the case of $\mathscr{C}^1(\mathbb{R}^d)$-functions with first order derivatives being Hölder-continuous with any parameter and the process given by a stochastic flow generated by a so-called $C^0(\mathbb{R}^d, \mathbb{R}^d)$-semimartingale. In both cases the quadratic variation process is expressed in terms of the generalised quadratic covariation process $\langle f'(X), X \rangle_t$ introduced by the authors in an earlier paper [71] (see also paper by Fuhrman and Tessitore [34], where authors extend the notion of the generalised quadratic covariation further to the infinite-dimensional case and non-differentiable functions). In [33] Föllmer, Protter and Shiryayev consider the case of an absolutely continuous function f

with a locally square integrable derivative and X a 1-dimensional Brownian motion, for which a version of Itô's formula is derived with the finite variation part expressed again in terms of the quadratic covariation $\langle f'(B), B\rangle_t$. The multidimensional case (where f belongs to the Sobolev space $\mathbb{W}^{1,2}$) is treated in [32]. In [51] Kendall discusses semimartingale decomposition of $r(B)$, where r is the distance function of Brownian motion on a manifold. The problem tackled in [51] is similar to ours as r fails to be differentiable on a set of measure zero, called the cut-locus. It is proved in [51] that $r(B)$ is a semimartingale and its canonical decomposition is found explicitly in the sequel [23].

In the proof of our main Theorem 1.2 we need the results of Carlen and Protter [19] and Barlow and Protter [4], which we discuss in the next section.

1.2 Convergence of semimartingales and convex functions

Paper of Barlow and Protter concerns convergence of semimartingales together with their semimartingale decomposition. We will present their results in the continuous setting, even though the original paper considered a wider class of semimartingales with jumps.

Suppose $\{X^n\}_{n\geq 1}$ is a sequence of continuous semimartingales with the decomposition $X^n = X_0^n + M^n + A^n$, such that $\lim_{n\to\infty} \mathbb{E}[(X^n - X)^*] = 0$. Here and in what follows we denote $X^* = \sup_t |X_t|$ and $X_t^* = \sup_{s\leq t} |X_s|$. Barlow and Protter prove that under some regularity conditions imposed on M^n and A^n not only that the limiting process X is again a continuous semimartingale, but that there is also convergence of the corresponding martingale and finite variation process parts of the decompositions.

Theorem 1.3. *([4, Thm. 1]) Let $\{X^n\}_{n\geq 1}$ be a sequence of continuous semimartingales in $\mathscr{H}^1$ with canonical decomposition $X^n = X_0^n + M^n + A^n$, satisfying for some constant K*

$$\mathbb{E}\Big[\int_0^\infty |dA_s^n|\Big] \leq K\ , \tag{1.1a}$$

$$\mathbb{E}[(M^n)^*] \leq K\ , \tag{1.1b}$$

and let X be a continuous process such that

$$\lim_{n\to\infty} \mathbb{E}[(X^n - X)^*] = 0 . \tag{1.2}$$

Then X is a continuous semimartingale in $\mathscr{H}^1$ and, moreover, if $X = X_0 + M + A$ is its decomposition, then

$$\mathbb{E}[M^*] \le K, \quad \mathbb{E}[\int_0^\infty |dA_s|] \le K$$

and

$$\lim_{n\to\infty} \|M^n - M\|_{\mathscr{H}^1} = 0 , \tag{1.3a}$$

$$\lim_{n\to\infty} \mathbb{E}[(A^n - A)^*] = 0 . \tag{1.3b}$$

In [19] Carlen and Protter prove that the assumptions of [4, Thm. 1] are satisfied in case when the sequence of $\mathscr{C}^2$ convex functions $\{f_n\}_{n\ge1}$ of (a not necessarily continuous) semimartingale $X = M + A$ converges to a convex f, thus making the result important in our situation.

We equip the set of convex functions on $\mathbb{R}^d$ with uniform convergence on compact sets with the corresponding metric ρ, defined by $\rho(f,g) = \sum_{k=1}^\infty 2^{-k}\rho_k(f,g)$, where

$$\rho_k(f,g) = \frac{\sup_{|x|\le k} |f(x) - g(x)|}{1 + \sup_{|x|\le k} |f(x) - g(x)|} .$$

Theorem 1.4. *([19, Thm. 1]) Let X be a continuous $\mathbb{R}^d$-valued semimartingale in $\mathscr{H}^1$ with $X_0 = 0$ and the decomposition $X = M + A$. For each $\alpha > 0$, define*

$$T_\alpha = \inf\{t > 0 : |X_t| > \alpha\} .$$

Let $\{f_n\}_{n\ge1}$ be a sequence of $\mathscr{C}^2$ convex functions on $\mathbb{R}^d$ and let f be a convex function on $\mathbb{R}^d$ such that $\lim_{n\to\infty} \rho(f_n, f) = 0$. Then $f(X_t)$ is a continuous semimartingale *and, moreover, if its decomposition is given by* $f(X_t) = f(X_0) + N_t + S_t$*, then for all* $\alpha > 0$

$$\lim_{n\to\infty} \|(N^n - N)^{T_\alpha}\|_{\mathscr{H}^1} = 0 \ ,$$
$$\lim_{n\to\infty} \mathbb{E}[(S^n - S)^{T_\alpha}] = 0 \ ,$$

where

$$N_t^n = \int_0^t \nabla f_n(X_s) dM_s$$

and

$$S_t^n = \int_0^t \nabla f_n(X_s) dA_s + \frac{1}{2} \sum_{i,j} \int_0^t \frac{\partial^2}{\partial x_i \partial x_j} f(X_s) d\langle X_s^i, X_s^j \rangle$$

are the martingale and finite variation parts of the decomposition of $f_n(X)$ *respectively.*

1.3 Elements of convex analysis

In order to prove the main result of this chapter, we require some notation and results from convex analysis. Proofs of the results stated in this section and more details on convex functions are given in [66]. See also [36]. We start by introducing some elementary notation.

1.3.1 Convex functions: some notation and results

Let f be any function living on $\mathbb{R}^d$ and taking values in $[-\infty, +\infty]$. At any point $x \in \mathbb{R}^d$ we define the *one-directional derivative of* f *with respect to a vector* $y \in \mathbb{R}^d$, if it exists, as follows

$$Df(x)[y] := \lim_{\lambda \downarrow 0} \frac{f(x + \lambda y) - f(x)}{\lambda} \ .$$

The two sided derivative at x in direction y exists if and only if $-Df(x)[-y]$, given by,

$$-Df(x)[-y] := \lim_{\lambda \uparrow 0} \frac{f(x + \lambda y) - f(x)}{\lambda}$$

is also well-defined and

$$Df(x)[y] = -Df(x)[-y] \tag{1.4}$$

Now, if the function f is convex, then the one-directional derivative always exists and, moreover, we may write

$$Df(x)[y] = \inf_{\lambda>0} \frac{f(x+\lambda y) - f(x)}{\lambda} \; . \tag{1.5}$$

Furthermore $Df(x)[y]$ is positively homogeneous (i.e. $Df(x)[\lambda y] = \lambda Df(x)[y]$ for $\lambda \in (0,\infty)$), convex in y with $Df(x)[0] = 0$ [66, Thm. 23.1] and

$$Df(x)[y] \geq -Df(x)[-y] \; . \tag{1.6}$$

If for a convex f defined on $\mathbb{R}^d$ and finite at some $x \in \mathbb{R}^d$ all directional derivatives at x exist, are two-sided and finite then we have ([66, Thm. 25.2])

$$Df(x)[y] = \langle \nabla f(x), y \rangle, \quad \forall y \in \mathbb{R}^d \; ,$$

where

$$\nabla f(x) := \left(\frac{\partial f}{\partial x_1}(x), \ldots, \frac{\partial f}{\partial x_d}(x) \right)$$

is the *gradient* of f at $x = (x_1, \ldots, x_d)$. Note that $\frac{\partial f}{\partial x_i}(x) = Df(x)[e_i]$, e_i being the i^{th} canonical basis vector of $\mathbb{R}^d$.

Of course a general convex function f is not necessarily everywhere differentiable, a simple example being $f(x) = |x|$ which is not differentiable at $x = 0$. We can however define a set of *subgradients* at each "troublesome" point like this.

Definition 1.5. *Let $f : \mathbb{R}^d \to \mathbb{R}$ be a convex function. A subgradient $\overline{\nabla} f(x)$ of f at $x \in \mathbb{R}^d$ is a gradient of an affine hyperplane $h(x) = \alpha + \beta^T x$, for $\alpha, \beta \in \mathbb{R}^d$, passing through the point $(x, f(x))$ and satisfying*

$$h(x') \leq f(x')$$

for all $x' \neq x$. We denote any subgradient at x in the direction of $y \in \mathbb{R}^n$ by $\langle \overline{\nabla} f(x), y \rangle$.

We say $h(x)$ is a *supporting hyperplane* of f at point $(x, f(x))$. Clearly at differentiable points $h(x)$ is unique and is just the tangent of f. Conversely, at points where f is not differentiable we can construct infinitely many tangential hyperplanes $h(x)$. The set of all subgradients at x is called the *subdifferential* of f at x, denoted $\partial f(x)$. A convex function with finite values is *subdifferentiable* everywhere. In subsequent sections we will need the following result

Theorem 1.6. *([66, Thm. 23.2]) Let f be a convex function and x a point at which f is finite. Then $\overline{\nabla} f(x)$ is a subgradient of f at x if and only if*

$$Df(x)[y] \geq \langle \overline{\nabla} f(x), y \rangle \qquad \forall y \in \mathbb{R}^d \backslash \{0\} \ . \tag{1.7}$$

The above says that a subgradient at x in the direction of y will always be less or equal to the one-sided directional derivative at x with respect to y. Relation (1.7) is called the *subgradient inequality* and can be used as an alternative definition of a subgradient.

Finally we mention the Lipschitz continuity property of convex functions.

Theorem 1.7. *([36, Ch. 3.1, Thm. 10]) Let f be a convex function on $\mathbb{R}^d$ and let U be an open convex subset of $\mathbb{R}^d$. Then f is continuous on U if and only if f is locally Lipschitz on U, i.e. for all $u \in U$ there exist constants $K > 0$ and $\epsilon > 0$ such that*

$$|f(x) - f(y)| \leq K\|x - y\|, \quad \forall x, y \in B_u(\epsilon) \ ,$$

where $B_u(\epsilon)$ is an open ball of radius ϵ centered at u and $\|\cdot\|$ is the usual Euclidean norm.

1.3.2 Differential theory of convex functions

This section is devoted to studying the set $\mathscr{D}$ of points in the domain of f at which the supporting hyperplane is unique. In what follows we assume that f is *proper*, i.e. $f(x) < +\infty$ for at least one x and $f(x) > -\infty$ for all x. By $\mathrm{dom} f$ we denote the *effective domain* of f, that is $\mathrm{dom} f = \{x \in \mathbb{R}^d : f(x) < \infty\}$. We denote by $\mathrm{int}(\mathrm{dom} f)$ the interior of $\mathrm{dom} f$.

Theorem 1.8. *([66, Thm. 25.2], [66, Thm. 25.4]) Let f be a proper convex function on $\mathbb{R}^d$, and let $\mathscr{D}$ be the set of points in $\mathrm{int}(\mathrm{dom} f)$ at which f is differentiable. Then*

$Df(x)[\cdot]$ is a linear function on $\mathbb{R}^d$ iff $x \in \mathscr{D}$. Moreover, $\mathscr{D}$ is dense in int(dom$f$), and its complement in int(dom$f$) is a set of measure zero.

In fact, in order for $Df(x)[\cdot]$ to be linear, and hence for f to be differentiable at x, it suffices that the partial derivatives with respect to the basis vectors of $\mathbb{R}^d$ exist at x.

So, we see that set $\mathscr{D}^c$ of points at which a convex function f does not have a unique supporting hyperplane has measure zero. Consequently any process which has a probability density at each time t spends time of measure zero in $\mathscr{D}^c$, an important fact we will use in the sequel.

To prove Theorem 1.2 for a general (continuous and proper but not necessarily differentiable) convex f we will approximate it by a sequence of twice continuously differentiable convex functions $f_n : \mathbb{R}^d \to \mathbb{R}$, to which we know Itô's formula can be applied. On top of this, working with convex functions gives us an advantage of being able to deduce from the pointwise convergence of the functions something about the convergence of their corresponding gradients.

Theorem 1.9. *(variation of [66, Thm. 25.7]) Let f be a convex function defined on $\mathbb{R}^d$ and $\{f_n\}$ a sequence of smooth convex functions on $\mathbb{R}^d$ such that $\lim_{n\to\infty} f_n(x) = f(x)$ $\forall x \in \mathbb{R}^d$. Let $\mathscr{D} \subseteq$ int(dom$f$) be the set of points where f is differentiable. Then*

$$\lim_{n\to\infty} \nabla f_n(x) = \nabla f(x) \qquad \forall x \in \mathscr{D} \, . \tag{1.8}$$

Proof. See proof of [66, Thm. 25.7]. □

This result will be used several times in Sections 5 and 6.

We conclude this section with a result concerning convergence of subgradients. Consider a sequence $\{x_n\}_{n\leq 1}$ with $x_n \in$ int(dom$f$), $n \geq 1$, and $x \in$ int(dom$f$) such that $\lim_{n\to\infty} x_n = x$. Of course in general $\lim_{n\to\infty} \overline{\nabla} f(x_n)$ need not exist. However the situation when $x_n = x + \epsilon_n y$ for some $y \in \mathbb{R}^d$ and $\epsilon_n \to 0$ as $n \to \infty$, i.e. when x_n approaches x from a single direction y, is special. In this case it is known that $\overline{\nabla} f(x_n)$ converges to the part of the boundary of $\partial f(x)$ consisting of points at which y is normal to $\partial f(x)$ [66, Thm. 24.6]. Moreover,

Proposition 1.10. *Let $f : \mathbb{R}^d \to \mathbb{R}$ be a convex function. For any $x \in \mathbb{R}^d$, for almost all $y \in S^{d-1}$, where S^{d-1} is the unit sphere in $\mathbb{R}^d$,*

$$\lim_{\epsilon \downarrow 0} \overline{\nabla} f(x + \epsilon y)$$

exists, belongs to $\partial f(x)$ and is unique for any selection $\overline{\nabla} f(x + \epsilon y) \in \partial f(x + \epsilon y)$ we may make from the subdifferential of f at $x + \epsilon y$ for any $\epsilon > 0$.

Proof. First of all recall that $Df(x)[y] = \lim_{\epsilon \downarrow 0}(f(x + \epsilon y) - f(x))/\epsilon$ is a positively homogeneous function, convex in y with $Df(x)[0] = 0$. Let $g(y) := Df(x)[y]$. Hence $\nabla g(\lambda y)$ exists and is unique for all $\lambda > 0$ for almost all $y \in \mathbb{R}^d$. Fix $x, y \in \mathbb{R}^d$ and without loss of generality, by adding a suitable affine function to f, assume that

$$f(x) = g(y) = \nabla g(y) = 0 \; .$$

We argue by contradiction. If theorem fails then we can find a subsequence $\epsilon_n \to 0$ and a selection $\overline{\nabla} f(x + \epsilon_n y) \in \partial f(x + \epsilon_n y)$ such that

$$\lim_{n \to \infty} \overline{\nabla} f(x + \epsilon_n y) = h \neq 0 \; , \tag{1.9}$$

and also a vector $u \in \mathbb{R}^d$ with $\langle h, u \rangle > 0$. For such u consider

$$\frac{f(x + \epsilon_n y + \epsilon_n \lambda u) - f(x + \epsilon_n y)}{\epsilon_n} = \lambda \frac{f(x + \epsilon_n y + \epsilon_n \lambda u) - f(x + \epsilon_n y)}{\epsilon_n \lambda} \; .$$

Using (1.5) and homogeneity of $g(y)$ the above is greater or equal to

$$\frac{\lambda}{\epsilon_n} Df(x + \epsilon_n y)[\epsilon_n u] = \lambda Df(x + \epsilon_n y)[u] \geq \lambda \langle \overline{\nabla} f(x + \epsilon_n y), u \rangle = \lambda \langle h, u \rangle + o(1)$$

where the last two inequality signs come from expressions (1.7) and (1.9) respectively, and $o(1) \to 0$ as $n \to \infty$. Thus we obtain

$$\frac{f(x + \epsilon_n y + \epsilon_n \lambda u) - f(x + \epsilon_n y)}{\epsilon_n} \geq \lambda \langle h, u \rangle + o(1) \; , \tag{1.10}$$

On the other hand, since $f(x) = g(y) = 0$, we have

$$\frac{f(x+\epsilon_n y) - f(x)}{\epsilon_n} = \frac{f(x+\epsilon_n y)}{\epsilon_n} = o(1) \,. \tag{1.11}$$

Hence combining (1.10) and (1.11) one obtains

$$\begin{aligned}\frac{f(x+\epsilon_n y+\epsilon_n \lambda u) - f(x)}{\epsilon_n} &= \frac{f(x+\epsilon_n y+\epsilon_n \lambda u) - f(x+\epsilon_n y)}{\epsilon_n} + \frac{f(x+\epsilon_n y) - f(x)}{\epsilon_n} \\ &\geq \lambda\langle h,u\rangle + o(1) \,.\end{aligned}$$

Letting $n \to \infty$, i.e. $\epsilon_n \to 0$, the above inequality becomes

$$\begin{aligned} & Df(x)[y+\lambda u] = g(y+\lambda u) \geq \lambda\langle h,u\rangle > 0 \\ \Rightarrow \quad & \frac{g(y+\lambda u)}{\lambda} = \frac{g(y+\lambda u) - g(y)}{\lambda} \geq \langle h,u\rangle > 0 \,.\end{aligned}$$

And so letting $\lambda \to 0$ one obtains

$$\langle \nabla g(y), u\rangle \geq \langle h,u\rangle > 0 \,.$$

But this contradicts the assumption that $\nabla g(y) = 0$.

$\square$

1.4 Piecewise linear convex functions and Meyer-Tanaka formula

In this section we start our analysis of the martingale part of $f(X)$. However, instead of treating the case of a general convex f we first prove Theorem 1.2 in a special case when f is piecewise linear. Using the Meyer-Tanaka formula, we will verify that any piecewise linear convex function of a continuous semimartingale is itself a continuous semimartingale and find the martingale part of the decomposition explicitly.

This result, although not essential, is a nice warm-up before we start dealing with a more general situation in the next sections. We refer reader to [65, Ch. VI.1] for a detailed discussion of classical Tanaka and Itô-Tanaka formulas (for $d = 1$). One might also find a discussion of convex functions in [65, Appendix.§3] useful.

Proposition 1.11. *Let $X = (X^1, ..., X^d)$ be a continuous semimartingale living on $\mathbb{R}^d$, with i^{th} component having decomposition $X_t^i = X_0^i + M_t^i + A_t^i$, $i \in \{1, ..., d\}$. Let $f : \mathbb{R}^d \to \mathbb{R}$ be a function defined by $f(x) = l_1(x) \vee ... \vee l_k(x)$, $x \in \mathbb{R}^d$, where $l_i(x) = \alpha_i + \sum_{j=1}^d \beta_{ij} x_j = \alpha_i + \beta_i^T x$, for $\alpha_i, \beta_i \in \mathbb{R}^d$, $i \in \{1, ..., k\}$, and $x \vee y := \sup\{x, y\}$. Then $f(X)$ is a semimartingale with decomposition*

$$f(X_t) = f(X_0) + \sum_{i=1}^{k} \int_0^t \mathbf{1}_{B_i}(X_s) \beta_i^T dX_s + \frac{1}{2} L_t \ , \tag{1.12}$$

where $B_i = \{x : \min\{k : \sup_j \{l_j(x)\} = l_k(x)\} = i\}$ and L_t is an increasing process, constant on the complement of $\{t : l_i(X_t) = l_j(X_t)$ for any $i \neq j\}$. In particular the local martingale part of $f(X)$ is given by

$$\sum_{i=1}^{k} \int_0^t \mathbf{1}_{B_i}(X_s) \beta_i^T dM_s \ . \tag{1.13}$$

Proof. We prove the proposition for the case when $k = 2$ and any $d \geq 1$ and general case follows by induction. Consider $f(x) = l_1(x) \vee l_2(x)$. Denote $l_1(X_t) = Y_t$ and $l_2(X_t) = Z_t$. Since X_t is a continuous semimartingale so are affine functionals, Y_t and Z_t, of X_t. Let the corresponding decompositions be $Y = M + A$ and $Z = N + S$. Consider $f(x) = l_1(x) \vee l_2(x) = y \vee z$. We can rewrite $y \vee z$ as follows

$$y \vee z = \frac{1}{2} \left(|y - z| + y + z \right) \ .$$

Hence, using the differential notation for simplicity, we obtain

$$d(Y_t \vee Z_t) = \frac{1}{2} d \left(|Y_t - Z_t| + Y_t + Z_t \right) = \frac{1}{2} \left(d(|W_t|) + dY_t + dZ_t \right) \ ,$$

where $W := Y - Z$, and so $W = (M - N) + (A - S)$. Using Meyer-Tanaka formula the above becomes

$$\frac{1}{2}\left(\operatorname{sgn}(W_t)dW_t + dL_t^0 + dY_t + dZ_t\right) ,$$

where L_t^0 is the local time of W at 0. Next

$$\begin{aligned}\frac{1}{2}\left(\operatorname{sgn}(W_t)d(M_t - N_t) + \operatorname{sgn}(W_t)d(A_t - S_t) + d(M_t + A_t) + d(N_t + S_t) + dL_t^0\right) = \\ = \frac{1}{2}\left[\left(\operatorname{sgn}(W_t)+1\right)dM_t - \left(\operatorname{sgn}(W_t)-1\right)dN_t + \right. \\ \left. + \left(\operatorname{sgn}(W_t)+1\right)dA_t - \left(\operatorname{sgn}(W_t)-1\right)dS_t + dL_t^0\right] .\end{aligned}$$

Now $\operatorname{sgn}(W_t) = \operatorname{sgn}(Y_t - Z_t) = \mathbf{1}_{[Y_t > Z_t]} - \mathbf{1}_{[Y_t \le Z_t]}$ and so $\operatorname{sgn}(W_t) + 1 = 2\mathbf{1}_{[Y_t > Z_t]}$ and $\operatorname{sgn}(W_t) - 1 = -2\mathbf{1}_{[Y_t \le Z_t]}$. Hence we obtain

$$\begin{aligned}d\left(Y_t \vee Z_t\right) = \mathbf{1}_{[Y_t > Z_t]}dM_t + \mathbf{1}_{[Y_t \le Z_t]}dN_t + \mathbf{1}_{[Y_t > Z_t]}dA_t + \mathbf{1}_{[Y_t \le Z_t]}dS_t + \frac{1}{2}dL_t = \\ = \mathbf{1}_{[Y_t > Z_t]}dY_t + \mathbf{1}_{[Y_t \le Z_t]}dZ_t + \frac{1}{2}dL_t\end{aligned}$$

or

$$Y_t \vee Z_t = Y_0 \vee Z_0 + \int_0^t \mathbf{1}_{[Y_s > Z_s]}dY_s + \int_0^t \mathbf{1}_{[Y_s \le Z_s]}dZ_s + \frac{1}{2}dL_t ,$$

where L_t is a continuous increasing process, constant on the complement of $\{t : l_1(X_t) = l_2(X_t)\}$. The above expression is exactly (1.12) for $n = 2$. Noticing that $x \vee y \vee z = (x \vee y) \vee z$, the general case follows by induction.

□

Clearly the integrand in (1.13) is a measurable selection of the multivalued map $\partial f(x)$ and so Theorem 1.2 holds in the special case of convex piecewise linear functions. To illustrate this result we consider our simple example again: for $f(x) = |x|$ $d = 1$, $k = 2$ and $l_1(x) = -x$ and $l_2(x) = x$ and so $B_1 = \{x : x < 0\}$, $B_2 = \{x : x \ge 0\}$ and L_t is an increasing process constant on the complement of $\{t : X_t = 0\}$.

1.5 Semimartingale decomposition of $f(\widetilde{X}_t)$

We are now ready to start the analysis of the general case of a convex f defined over the whole of Euclidean space $\mathbb{R}^d$. Let X be a continuous semimartingale in $\mathbb{R}^d$ with decomposition $X = M + A$ and defined on some filtered probability space $(\Omega, \{\mathscr{F}_t\}_{t\geq 0}, \mathscr{F}, \mathbb{P})$. Let $(\widetilde{\Omega}, \{\widetilde{\mathscr{F}}_t\}_{t\geq 0}, \widetilde{\mathscr{F}}, \widetilde{\mathbb{P}})$ be some enlargement of this space and let B be an $(\widetilde{\mathscr{F}}_t)$-standard Brownian motion independent of X. Define the *perturbed process* $\widetilde{X}$ on $(\widetilde{\Omega}, \{\widetilde{\mathscr{F}}_t\}_{t\geq 0}, \widetilde{\mathscr{F}}, \widetilde{\mathbb{P}})$ by

$$\widetilde{X}^{(\epsilon)}_t := \widetilde{X}_t := X_t + \epsilon B_t, \qquad \epsilon > 0, \quad t \geq 0 \ .$$

For simplicity of notation we shall suppress the superscript (ϵ) wherever possible. For simplicity also but without loss of generality we can assume that $X_0 = \widetilde{X}_0 = 0$.

In this section we find the martingale part of $f(\widetilde{X}^{(\epsilon)})$ explicitly in order to take the limit as $\epsilon \to 0$ in the next section and hence prove Theorem 1.2. The reasoning behind adding a small amount of Brownian motion to X_t is as follows: we know very little about the behaviour of X_t as it is a general semimartingale. For instance, it can at some times be trivial, i.e. constant. Hence it might spend positive amount of time in those points where f is not differentiable, that is, where it has more than one supporting hyperplane. To avoid this happening we perturb X_t by adding ϵB_t. Then

Lemma 1.12. *$\widetilde{X}_t$ has a probability density at each $t > 0$ and, in particular, spends zero time in any null set.*

Proof. It suffices to prove that $\widetilde{\mathbb{P}}(\widetilde{X}_t \in N) = 0$ for any $t > 0$ and $N \subset \mathbb{R}^d$ with $Leb(N) = 0$. Then it will follow that for all $t > 0$ the law of $\widetilde{X}_t$ under $\widetilde{\mathbb{P}}$ is absolutely continuous with respect to the Lebesgue measure, and the corresponding Radon-Nikodym derivative is the probability density of $\widetilde{X}_t$. For any Lebesgue-null set N we have

$$\widetilde{\mathbb{P}}(\widetilde{X}_t \in N) = \mathbb{E}\left[\widetilde{\mathbb{P}}(\widetilde{X}_t \in N | \mathscr{F}_t)\right] \ ,$$

where $\mathscr{F}_t = \sigma(\{X_s; 0 \leq s \leq t\})$, and we use the tower property of conditional expectation. Next we express $\widetilde{X}_t$ in terms of X_t and B_t and use the fact that B_t is independent of X_t, and hence of $\mathscr{F}_t$, to obtain

$$\mathbb{E}\big[\widetilde{\mathbb{P}}(X_t + \epsilon B_t \in N | \mathscr{F}_t)\big] = \int \widetilde{\mathbb{P}}(x + \epsilon B_t \in N) d\mu_t(x) ,$$

where μ_t is the law of X_t. Observe that $\hat{B}_t := x + \epsilon B_t$ is a Brownian motion started at x with $\langle \hat{B}_t, \hat{B}_t \rangle = \epsilon^2 t$. But we know that Brownian motion hits null-sets with probability zero. Hence, the above integral is equal to zero and the lemma is proved.

□

In Section 1.3.2 we have seen that $\mathscr{D}^c$, the set of points at which f is not differentiable, is Lebesgue-null. Consequently, by the above lemma, $\widetilde{X}$ spends zero time at those "ambiguous" points. Hence, $\nabla f(\widetilde{X})$ is almost surely everywhere defined. Moreover, because of Lemma 1.12 a particular measurable choice of $\overline{\nabla} f(x) \in \partial f(x)$ at $x \in \mathscr{D}^c$ is unimportant as it does not change the value of the stochastic integral $\int_0^t \overline{\nabla} f(\widetilde{X}_s) d\widetilde{M}_s$, which we will show is the martingale part of $f(\widetilde{X})$. To do that we approximate f by a sequence of convex *twice continuously differentiable* functions.

Let $\{f_n\}_{n\geq 1}$ be a sequence of twice continuously differentiable convex functions on $\mathbb{R}^d$ converging to f with respect to ρ, i.e. such that $\lim_{n\to\infty} \rho(f_n, f) = 0$, where metric ρ is as defined in Section 1.2. We need to prove that stochastic integral $\int_0^t \nabla f_n(\widetilde{X}_s) d\widetilde{M}_s$, the martingale part of $f_n(\widetilde{X})$, converges in some sense to $\int_0^t \overline{\nabla} f(\widetilde{X}_s) d\widetilde{M}_s$ for some measurable choice of $\overline{\nabla} f(x) \in \partial f(x)$, and that it is indeed the martingale part of $f(\widetilde{X})$.

We need to note the following two inequalities

Lemma 1.13. *(Adapted from [19, Lemma, p. 2])*

$$\sup_n \sup_{|x|\leq r} |\nabla f_n(x)| \leq C_r < \infty, \quad \forall r > 0 , \tag{1.14}$$

and

$$\sup_{|x|\leq r} |\overline{\nabla} f(x)| \leq C_r < \infty, \quad \forall r > 0 , \tag{1.15}$$

where C_r is some constant only depending on r and $\overline{\nabla} f(x)$ is any choice of subgradient

$\partial f(x)$.

Proof. To see why inequality (1.14) is true, first notice that, since $\lim_{n\to\infty} \rho(f_n, f) = 0$, the variation of the convex functions f_n is uniformly bounded in n on $\{|x| \le r+1\}$ for any $r > 0$. Denote this bound by C_r. Let x_n be such that

$$\nabla f_n(x_n) = \sup_{|x|\le r} |\nabla f_n(x)|$$

and let $u_n := \nabla f_n(x_n)/|\nabla f_n(x_n)|$. Then

$$\begin{aligned} |\nabla f_n(x_n)| = \langle \nabla f_n(x_n), \frac{\nabla f_n(x_n)}{|\nabla f_n(x_n)|}\rangle = \langle \nabla f_n(x), u_n\rangle = Df_n(x)[u_n] \\ = \inf_{\lambda>0} \frac{f_n(x_n + \lambda u_n) - f_n(x_n)}{\lambda} \le f_n(x_n + u_n) - f_n(x_n)\,. \end{aligned} \tag{1.16}$$

But, since $|x_n + u_n| \le r+1$, the above is less than C_r for all n and (1.14) follows.

Now, since f_n converges to f uniformly on compact sets, we also have $f_n \to f$ pointwise. Therefore, for any x, y with $|x|, |y| < r+1$ the inequality $f_n(x) - f_n(y) \le C_r$, $\forall n \ge 1$, (which follows since C_r bounds the variation of f_n's) implies $f(x) - f(y) \le C_r$ by virtue of taking the limit $n \to \infty$. So, by a calculation similar to (1.16), we have for any $\overline{\nabla} f(x) \in \partial f(x)$

$$|\overline{\nabla} f(x^*)| = \langle \overline{\nabla} f(x^*), u^*\rangle \le Df(x^*)[u^*] \le f(x^* + u^*) - f(x^*) \le C_r\ ,$$

where x^* is such that $\overline{\nabla} f(x^*) = \sup_{|x|\le r} |\overline{\nabla} f(x)|$ and $u^* := \overline{\nabla} f(x^*)/|\overline{\nabla} f(x^*)|$ for any $r > 0$. □

We are now ready to prove the following

Lemma 1.14. *The local martingale part of $f(\widetilde{X}_t)$ is given by the limit*

$$\lim_{n\to\infty} \int_0^t \nabla f_n(\widetilde{X}_s) d\widetilde{M}_s = \int_0^t \overline{\nabla} f(\widetilde{X}_s) d\widetilde{M}_s \tag{1.17}$$

locally in $\mathscr{H}^1$, where $\overline{\nabla} f(x) \in \partial f(x)$ is some measurable choice of subgradient of f at x.

Proof. Since for each $n \geq 1$ f_n is in $\mathscr{C}^2$, the martingale part of $f_n(\widetilde{X})$ is given by $\int \nabla f_n(\widetilde{X}) d\widetilde{M}$, where $\widetilde{M} = M + \epsilon B$. Theorem 1.4, applied to our sequence $\{f_n\}_{n\geq 1}$ and the semimartingale $\widetilde{X}$, then ensures that the martingale part of the limiting process $f(\widetilde{X}_t)$ is given by the limit of $\int \nabla f_n(\widetilde{X}) d\widetilde{M}$ as n tends to infinity, locally in $\mathscr{H}^1$. Our aim is to prove that this limit is indeed equal to $\int \overline{\nabla} f(\widetilde{X}) d\widetilde{M}$ for some measurable choice of subgradient $\overline{\nabla} f \in \partial f$.

Let $B(r)$ be an open ball of radius r and $B(r')$ an open ball of radius r' with $r' > r > 0$, both centred at the origin. For all $r, r' > 0$ define stopping times $T_r := \inf\{t : X_t \notin B(r)\}$ and $\widetilde{T}_{r'} := \inf\{t : \widetilde{X}_t \notin B(r')\}$ and take $\widetilde{T} = T_r \wedge \widetilde{T}_{r'}$. Assume also that $\widetilde{X}_{t\wedge\widetilde{T}}, X_{t\wedge\widetilde{T}} \in \mathscr{H}^1$ for all $t \geq 0$; we know that continuous semimartingales are at least locally in $\mathscr{H}^1$. We consider the stopped process $\widetilde{X}_{t\wedge\widetilde{T}}$. Note that $X_{t\wedge\widetilde{T}} \in B(r) \subset B(r')$ and $\widetilde{X}_{t\wedge\widetilde{T}} \in B(r')$ for all $t \geq 0$. By Lemma 1.12 the law of $\widetilde{X}_{t\wedge\widetilde{T}}$ has a density for all $t < \widetilde{T}$.

Note that for proving Lemma 1.14 it would have sufficed to stop $\widetilde{X}$ at $\widetilde{T}_{r'}$. However, in order to be consistent with localisation we will be using to prove Theorem 1.2, we use $\widetilde{T} = T_r \wedge \widetilde{T}_{r'}$ instead.

Notice that convergence of a continuous (local) martingale M in $\mathscr{H}^p$ is equivalent to convergence of $\langle M, M\rangle^{1/2}$ in $\mathscr{L}^p$. So, in this case convergence in $\mathscr{H}^p$ implies convergence in $\mathscr{H}^l$ for $1 \leq l < p$. In our case it is easier to prove convergence (1.17) in $\mathscr{H}^2$ and then deduce convergence in $\mathscr{H}^1$. For any measurable selection $\overline{\nabla} f \in \partial f$ and $t > 0$ we have

$$
\begin{aligned}
&\lim_{n\to\infty} \Big|\Big| \int_0^{t\wedge\widetilde{T}} \left(\nabla f_n(\widetilde{X}_s) - \overline{\nabla} f(\widetilde{X}_s)\right) d\widetilde{M}_s \Big|\Big|_{\mathscr{H}^2} \\
&= \lim_{n\to\infty} \mathbb{E}\Big[\int_0^{t\wedge\widetilde{T}} \left(\nabla f_n(\widetilde{X}_s) - \overline{\nabla} f(\widetilde{X}_s)\right)^2 d\langle \widetilde{M}, \widetilde{M}\rangle_s \Big]^{1/2} .
\end{aligned}
$$

Using inequalities (1.14) and (1.15) we can bound the expression inside the expectation sign above as follows

$$\int_0^{t\wedge\widetilde{T}} \left(\nabla f_n(\widetilde{X}_s) - \overline{\nabla} f(\widetilde{X}_s)\right)^2 d\langle \widetilde{M}, \widetilde{M}\rangle_s \leq 4C_{r'}^2 \int_0^{t\wedge\widetilde{T}} d\langle \widetilde{M}, \widetilde{M}\rangle_s$$
$$\leq 4C_{r'}^2 \langle \widetilde{M}, \widetilde{M}\rangle_{t\wedge\widetilde{T}} < \infty \,,$$

where the quadratic variation $\langle \widetilde{M}, \widetilde{M}\rangle_{t\wedge\widetilde{T}}$ is finite, because it is the bracket of a bounded continuous semimartingale $\widetilde{X}_{t\wedge\widetilde{T}}$ (see [65, Ch. IV, Thm. 1.3]). Using dominated convergence theorem we can now take the limit inside the expectation sign and, since the integrand is bounded above by $4C_{r'}^2$, we can also pull the limit inside the integral sign. We can then use almost sure convergence of $\nabla f_n(\widetilde{X}_t)$ to $\nabla f(\widetilde{X}_t)$ for all $\widetilde{X}_t \in \mathscr{D}$ and the fact that particular choices $\overline{\nabla} f(\widetilde{X}_t) \in \partial f(\widetilde{X}_t)$ for $\widetilde{X}_t \in \mathscr{D}^c$ are not charged by the integral to conclude that the limit in question is equal to

$$\mathbb{E}\Big[\int_0^{t\wedge\widetilde{T}} \lim_{n\to\infty} \left(\nabla f_n(\widetilde{X}_s) - \overline{\nabla} f(\widetilde{X}_s)\right)^2 d\langle \widetilde{M}, \widetilde{M}\rangle_s\Big]^{1/2} = 0 \,.$$

It follows that $\int_0^{t\wedge\widetilde{T}} \nabla f_n(\widetilde{X}_s) d\widetilde{M}_s$ converges to $\int_0^{t\wedge\widetilde{T}} \overline{\nabla} f(\widetilde{X}_s) d\widetilde{M}_s$ in $\mathscr{H}^2$, and hence in $\mathscr{H}^1$. This is true for any radiuses $r' > r > 0$ of localisation, and so (1.17) follows.

□

We also prove the following lemma concerning the semimartingale decomposition of $f(\widetilde{X})$ that we will require for the proof Theorem 1.2.

Lemma 1.15. *Let $\widetilde{N}_t^{(\epsilon)} = \int_0^t \overline{\nabla} f(\widetilde{X}_s^{(\epsilon)}) d\widetilde{M}_s^{(\epsilon)}$ and $\widetilde{S}^{(\epsilon)}$ be the martingale and the finite variation parts of the decomposition of $f(\widetilde{X}^{(\epsilon)})$ respectively. Then for all $\epsilon \leq 1$*

$$\mathbb{E}\Big[\sup_{t\leq\widetilde{T}} |\widetilde{N}_t^{(\epsilon)}|\Big] \leq K_{r,r'} \,, \tag{1.18a}$$

$$\mathbb{E}\Big[\int_0^{\widetilde{T}} |d\widetilde{S}_s^{(\epsilon)}|\Big] \leq K_{r,r'} \,, \tag{1.18b}$$

where $K_{r,r'}$ is a constant depending on r and r' and independent of ϵ.

Proof. The proof largely follows proof of Theorem 1.4 in [19]: we prove that the sequence of continuous semimartingales $\{f_n(\widetilde{X})\}_{n\geq 1}$ satisfies the conditions of Theorem 1.3 of Barlow and Protter, i.e.

$$\lim_{n\to\infty} \mathbb{E}\Big[\sup_{t\leq \widetilde{T}} |f_n(\widetilde{X}_t) - f(\widetilde{X}_t)|\Big] = 0 \tag{1.19a}$$

$$\sup_{n\geq 1} \mathbb{E}\Big[\sup_{t\leq \widetilde{T}} |\widetilde{N}^n_t|\Big] \leq K_{r,r'}\ , \tag{1.19b}$$

$$\sup_{n\geq 1} \mathbb{E}\Big[\int_0^{\widetilde{T}} |d\widetilde{S}^n_s|\Big] \leq K_{r,r'}\ , \tag{1.19c}$$

where for each $n \geq 1$ $\widetilde{N}^n$ and $\widetilde{S}^n$ are the martingale and the finite variation part of the decomposition of $f_n(\widetilde{X})$ respectively. Then (1.18a) and (1.18b) will follow by Theorem 1.3. The difference is only in the fact that we need to ensure that for small enough ϵ the constant $K_{r,r'}$ above can be taken to be independent of ϵ.

First of all notice that (1.19a) follows from the fact that $\lim_{n\to\infty} \rho(f_n, f) = 0$. Next consider (1.19b). By the Burkholder-Davis-Gundy inequality we have for some constant $p < \infty$

$$\begin{aligned}
\mathbb{E}\Big[\sup_{t\leq \widetilde{T}} |\widetilde{N}^n_t|\Big] \leq p\mathbb{E}\Big[\langle \widetilde{N}^n, \widetilde{N}^n\rangle^{1/2}_{\widetilde{T}}\Big] &= p\mathbb{E}\Big[\Big(\int_0^{\widetilde{T}} |\overline{\nabla} f_n(\widetilde{X}_t)|^2 d\langle \widetilde{M}, \widetilde{M}\rangle_t\Big)^{1/2}\Big] \\
&\leq pC_{r'}\mathbb{E}\Big[\langle \widetilde{M}, \widetilde{M}\rangle^{1/2}_{\widetilde{T}}\Big] \\
&= pC_{r'}\mathbb{E}\Big[(\langle M, M\rangle_{\widetilde{T}} + \epsilon^2 \widetilde{T})^{1/2}\Big]\ .
\end{aligned}$$

To finish we need to bound $\langle \widetilde{M}, \widetilde{M}\rangle_{\widetilde{T}}$ by some constant independent of ϵ. We have $\langle \widetilde{M}, \widetilde{M}\rangle_{\widetilde{T}} = \langle M, M\rangle_{\widetilde{T}} + \epsilon^2 \widetilde{T}$ which for all $\epsilon \leq 1$ is less than $\langle M, M\rangle_{\widetilde{T}} + \widetilde{T}$ which is in turn bounded above by $\langle M, M\rangle_{T_r} + T_r$, since $T_r \geq \widetilde{T} = T_r \wedge \widetilde{T}_{r'}$. Hence, eventually for all ϵ

$$\mathbb{E}\Big[\sup_{t\leq \widetilde{T}} |\widetilde{N}^n_t|\Big] \leq pC_{r'}\mathbb{E}\Big[(\langle M, M\rangle_{T_r} + T_r)^{1/2}\Big]\ ,$$

where the right-hand side is independent of ϵ as well as n, and so (1.19b) follows.

We now turn to (1.19c). The finite variation part of $f_n(\widetilde{X}_{t\wedge\widetilde{T}})$ is given by

$$\widetilde{S}^n_{t\wedge\widetilde{T}} = \int_0^{t\wedge\widetilde{T}} \nabla f_n(\widetilde{X}_s) dA_s + \frac{1}{2}\int_0^{t\wedge\widetilde{T}} \sum_{i,j} \frac{\partial^2 f_n}{\partial x_i \partial x_j}(\widetilde{X}_s) d\langle \widetilde{X}_s, \widetilde{X}_s\rangle \ . \tag{1.20}$$

Writing $Var_{s\le t}(Y_s)$ for $\int_0^t |dY_s|$ for a continuous process Y_t, we have

$$\begin{aligned}\int_0^{\widetilde{T}} |d\widetilde{S}^n_s| &= Var_{t\le\widetilde{T}}(\widetilde{S}^n_t)\\ &\le Var_{t\le\widetilde{T}}\Big(\int_0^t \nabla f_n(\widetilde{X}_s)dA_s\Big) + Var_{t\le\widetilde{T}}\Big(\frac{1}{2}\int_0^t \sum_{i,j}\frac{\partial^2 f_n}{\partial x_i\partial x_j}(\widetilde{X}_s)d\langle\widetilde{X}_s,\widetilde{X}_s\rangle\Big) \ .\end{aligned}$$

Now

$$Var_{t\le\widetilde{T}}\Big(\int_0^t \nabla f_n(\widetilde{X}_s)dA_s\Big) \le C_{r'} Var_{t\le\widetilde{T}}\Big(\int_0^t dA_s\Big) \le C_{r'}\int_0^{t\wedge T_r} |dA_s| \ , \tag{1.21}$$

where we have used $T_r \ge T_r \wedge \widetilde{T}_{r'} = \widetilde{T}$ and the fact that Var_t is an increasing process of t. Hence, the bound on the expectation of the RHS above depends on r and r' but not on ϵ.

Next we observe that, since f_n is convex and the bracket $\langle\widetilde{X}_s,\widetilde{X}_s\rangle$ is an increasing process of t, the second summand of (1.20), which we denote by F^n_t, is also an increasing function of t. Thus,

$$Var_{t\le\widetilde{T}}\Big(\frac{1}{2}\int_0^t \sum_{i,j}\frac{\partial^2 f_n}{\partial x_i\partial x_j}(\widetilde{X}_s)d\langle\widetilde{X}_s,\widetilde{X}_s\rangle\Big) \le F^n_{\widetilde{T}} \le F^n_{T_r} \ .$$

But, because f_n is twice continuously differentiable, using Itô's formula we can also write

$$\mathbb{E}[F^n_{T_r}] = \mathbb{E}[f_n(\widetilde{X}_{T_r}) - f_n(\widetilde{X}_0)] - \mathbb{E}\Big[\int_0^{T_r} \nabla f_n(\widetilde{X}_s)dA_s\Big] \ .$$

Using Lemma 1.13 and Lipschitz continuity of f_n in $B(r')$ (we write $K_{r'}$ for the corresponding Lipschitz constant) we have

$$\begin{aligned}\mathbb{E}[F^n_{T_r}] \le & K_{r'}\mathbb{E}[|\widetilde{X}_{T_r}|] + C_{r'}\mathbb{E}[|A_{T_r}|] \\ = & K_{r'}\mathbb{E}[|X_{T_r} + \epsilon B_{T_r}|] + C_{r'}\mathbb{E}[|A_{T_r}|] \\ \le & K_{r'}(r + \mathbb{E}[B_{T_r}]) + C_{r'}\mathbb{E}[|A_{T_r}|]\end{aligned}$$

for all $\epsilon \le 1$, i.e. the bound on $\mathbb{E}[F^n_{T_r}]$ is independent of both n and ϵ for all $\epsilon \le 1$. Combining this with (the expectation of) the bound (1.21) we obtain (1.19c).

The assertion of the lemma now follows by Theorem 1.3. □

1.6 Proof of Theorem 1.2

Finally we need to derive the analogous result for our original object of interest, a continuous semimartingale X.

Proof of Theorem 1.2. We have $\lim_{\epsilon\downarrow 0}\widetilde{X}^{(\epsilon)} = X$ almost surely and, thus, for a continuous convex f, $\lim_{\epsilon\to 0} f(\widetilde{X}^{(\epsilon)}) = f(X)$ almost surely. Note that the limit of the process $\widetilde{X}^{(\epsilon)}$ as ϵ tends to zero lives in the enlarged probability space $(\widetilde{\Omega}, \widetilde{\mathscr{F}}, \widetilde{\mathbb{P}})$, even though the original process X is defined on $(\Omega, \mathscr{F}, \mathbb{P})$. Crucially by Itô's lemma $f(\widetilde{X}^{(\epsilon)}_t)$ is a continuous semimartingale for every $\epsilon > 0$. Hence, we can apply Theorem 1.3 to the sequence of semimartingales $\{f(\widetilde{X}^{(\epsilon)})\}_{\epsilon>0}$ if we can show that the conditions (1.1a), (1.1b) and (1.2) of the theorem are satisfied in our case.

We use the same localisation as in the proof of Lemma 1.14, i.e. we consider $\widetilde{X}_{t\wedge\widetilde{T}}$ for $\widetilde{T} = T_r \wedge \widetilde{T}_{r'} = \inf\{t : X_t \notin B(r)\} \wedge \inf\{t : \widetilde{X}_t \notin B(r')\}$, with $r' > r > 0$.

In view of Lemma 1.15 the only thing we need to prove in order to apply results of Theorem 1.3 is

$$\lim_{\epsilon\downarrow 0}\mathbb{E}\Big[\sup_{t\le\widetilde{T}}|f(\widetilde{X}^{(\epsilon)}_t) - f(X_t)|\Big] = 0\,,$$

Since by Theorem 1.7 f is Lipschitz in the ball $B(r')$, we have

$$\mathbb{E}\Big[\sup_{t\le\widetilde{T}}|f(\widetilde{X}_t) - f(X_t)|\Big] \le K_{r'}\mathbb{E}\Big[\sup_{t\le\widetilde{T}}|\widetilde{X}_t - X_t|\Big] = \epsilon K_{r'}\mathbb{E}\Big[\sup_{t\le\widetilde{T}}|B_t|\Big]\,,$$

where $K_{r'} < \infty$ is a Lipschitz constant depending on r'. Taking the limit $\epsilon \to 0$ gives the desired result.

Together with expressions (1.18a) and (1.18b) of Lemma 1.15 this ensures that the conditions of Theorem 1.3 are satisfied in our case. It follows immediately that the martingale part of $f(X)$ is given by the limit as $\epsilon \to 0$ of $\int \overline{\nabla} f(\widetilde{X}^{(\epsilon)}) d\widetilde{M}^{(\epsilon)}$, the martingale part of $f(\widetilde{X}^{(\epsilon)})$, locally in $\mathscr{H}^1$. All is left to prove now is that this limit is given by $\int \overline{\nabla} f(X) dM$ for some choice of $\overline{\nabla} f(x) \in \partial f(x)$, i.e. that for all $t > 0$

$$\lim_{\epsilon \downarrow 0} \int_0^t \overline{\nabla} f(\widetilde{X}^{(\epsilon)}_{s\wedge\widetilde{T}}) d\widetilde{M}^{(\epsilon)}_{s\wedge\widetilde{T}} = \int_0^t \overline{\nabla} f(X_{s\wedge T_r}) dM_{s\wedge T_r} \tag{1.22}$$

in $\mathscr{H}^1$ for all $r' > r > 0$.

Proving the above convergence will require us to consider the limit of $\overline{\nabla} f(\widetilde{X}^{(\epsilon)}_{t\wedge\widetilde{T}})$ as ϵ tends to 0. From Proposition 1.10 we know that for all $t \geq 0$ for almost all values of B_t the limit $\lim_{\epsilon\downarrow 0} \overline{\nabla} f(X_t + \epsilon B_t)$ exists and belongs to $\partial f(X_t)$. Denote this limit by $\overline{\nabla} f(X_t)$. Also for any path of X and B for small enough ϵ, i.e. eventually for all ϵ, we have $T_r < \widetilde{T}_{r'}$. That is $\widetilde{T} \to T_r$ as $\epsilon \to 0$ a.s. and so

$$\lim_{\epsilon \downarrow 0} \overline{\nabla} f(X_{t\wedge\widetilde{T}} + \epsilon B_{t\wedge\widetilde{T}}) = \overline{\nabla} f(X_{t\wedge T_r}) \quad \text{a.s.}\,. \tag{1.23}$$

Again we consider convergence in $\mathscr{H}^2$ first, and convergence in $\mathscr{H}^1$ follows. We have, using the fact that $\lim_{\epsilon\downarrow 0} \widetilde{M}_{t\wedge\widetilde{T}} = \lim_{\epsilon\downarrow 0}(M_{t\wedge\widetilde{T}} + \epsilon B_{t\wedge\widetilde{T}}) = \lim_{\epsilon\downarrow 0} M_{t\wedge\widetilde{T}}$ a.s.

$$\begin{aligned}
\lim_{\epsilon \downarrow 0} \Big|\Big| \int_0^t \overline{\nabla} f(\widetilde{X}_{s\wedge\widetilde{T}}) d\widetilde{M}_{s\wedge\widetilde{T}} - \int_0^t \overline{\nabla} f(X_{s\wedge T_r}) dM_{s\wedge T_r} \Big|\Big|_{\mathscr{H}^2} \\
= \lim_{\epsilon\downarrow 0} \mathbb{E}\Big[\int_0^{\widetilde{T}} \overline{\nabla} f(\widetilde{X}_s)^2 d\langle M_s, M_s\rangle + \int_0^{T_r} \overline{\nabla} f(X_s)^2 d\langle M_s, M_s\rangle \\
- 2\int_0^\infty \overline{\nabla} f(\widetilde{X}_{s\wedge\widetilde{T}}) \overline{\nabla} f(X_{s\wedge T_r}) d\langle M_{s\wedge\widetilde{T}}, M_{s\wedge T_r}\rangle \Big]^{1/2}\,.
\end{aligned} \tag{1.24}$$

Once again we can use Lemma 1.13 to see that the first integrand in (1.24) is bounded above by $C^2_{r'} < \infty$, while the third integrand is bounded above by $C_r C_{r'} < \infty$.

Additionally we have

$$\int_0^{\widetilde{T}} \overline{\nabla} f(\widetilde{X}_s)^2 d\langle M_s, M_s\rangle \leq C_{r'}^2 \langle M_{\widetilde{T}}, M_{\widetilde{T}}\rangle < \infty$$

and

$$\int_0^{\infty} \overline{\nabla} f(\widetilde{X}_{s\wedge\widetilde{T}}) \overline{\nabla} f(X_{s\wedge T_r}) d\langle M_{s\wedge\widetilde{T}}, M_{s\wedge T_r}\rangle \leq C_r C_{r'} \langle M_{\widetilde{T}}, M_{\widetilde{T}}\rangle < \infty \ ,$$

since $\langle M_t, M_t\rangle$, is finite for any $t \leq \widetilde{T}$. Appealing to the dominated and bounded convergence theorems we can interchange the limit with the expectation and the integration signs respectively. Convergence (1.23) and the fact that $\widetilde{T} \to T_r$ a.s. then finally yield (1.22). Noticing that the above is true for all $r' > r > 0$ concludes the proof.

□

Example. As was mentioned before, in [16] Bouleau has proved that *any* measurable choice of subgradient $\overline{\nabla} f(X_t)$ works for the stochastic integral of Theorem 1.2. A function

$$\overline{\nabla}^e f(x) = \lim_{\theta\downarrow 0} \mathbb{E}[\overline{\nabla} f(x + \theta N)] \ , \tag{1.25}$$

where N is a standard d-dimensional Gaussian random variable, is a particular example. $\overline{\nabla}^e f(x)$ can be regarded as a sort of an average of (sub)gradients within the vicinity of x. To verify that it does indeed define a subgradient of f at each $x \in \mathbb{R}^d$ we check the subgradient inequality (1.7) of Theorem 1.6. For any $y \in \mathbb{R}^d \backslash \{0\}$ we have

$$\langle \overline{\nabla}^e f(x), y\rangle = \langle \lim_{\theta\downarrow 0} \mathbb{E}[\overline{\nabla} f(x + \theta N)], y\rangle = \lim_{\theta\downarrow 0} \mathbb{E}[\langle \overline{\nabla} f(x + \theta N), y\rangle] \ . \tag{1.26}$$

Now by the Lipschitz property of f and by the subgradient inequality (1.7) we have

$$\begin{aligned}\langle\overline{\nabla}f(x+\theta N),y\rangle \le D(x+\theta N)[y] &= \inf_{\lambda>0}\frac{f(x+\theta N+\lambda y)-f(x+\theta N)}{\lambda}\\ &\le f(x+\theta N+y)-f(x+\theta N)\le K|y|\end{aligned}$$

for some Lipshitz constant $K<\infty$ depending on x and N. Appealing to the bounded convergence theorem now allows us to take the limit inside the expectation in equation (1.26) above

$$\langle\overline{\nabla}^e f(x),y\rangle = \mathbb{E}[\langle\lim_{\theta\downarrow 0}\overline{\nabla}f(x+\theta N),y\rangle]\ . \tag{1.27}$$

But by Proposition 1.10 $\lim_{\theta\downarrow 0}\overline{\nabla}f(x+\theta N)$ exists, is unique and belongs to $\partial f(x)$ for almost all N. Denote this limit by $\overline{\nabla}^* f(x)$. Then (1.27) is equal to

$$\mathbb{E}[\langle\overline{\nabla}^* f(x),y\rangle]\le\mathbb{E}[Df(x)[y]]=Df(x)[y]\ .$$

Hence we have $\langle\overline{\nabla}^e f(x),y\rangle\le Df(x)[y]$ for any $y\in\mathbb{R}^d\backslash\{0\}$ for all x, and so $\overline{\nabla}^e f(x)$ is a well-defined subgradient of f.

Chapter 2

Elements of representation theory.

Starting with the present chapter the thesis is taking up a study of a different subject to that of Chapter 1. Chapters 3, 4 and 5 are concerned with a certain family of bivariate Markov chains, diffusions and repelling particle systems, respectively.

In this chapter we outline some basic results from representation theory which we will employ in later chapters. In particular we will describe finite-dimensional representations of the universal enveloping algebra $U(\mathfrak{gl}(n))$ and its quantisation $U_q(\mathfrak{gl}(n))$, treating separately the case of $U(\mathfrak{sl}(2))$ and $U_q(\mathfrak{sl}(2))$, and explain how they are related to combinatorics. We start with some basic definitions and notation.

2.1 Basics

Given a group G, Representation theory is concerned with finding all finite- and infinite-dimensional vector spaces on which elements of the group act as automorphisms and describing these actions. Thus homomorphisms, structure preserving maps between two algebraic objects, play a fundamental role in Representation theory. A *group homomorphism between two groups* $(G, \cdot)$ and $(H, \times)$ is a map $\varphi : G \to H$ such that $\varphi(g_1 \cdot g_2) = \varphi(g_1) \times \varphi(g_2)$. So, a group homomorphism preserves the identity and the inverses and so the structure of the group. An *isomorphism* is a homomorphism which is a bijection, an *endomorphism* is a homomorphism of an object into itself and an *automorphism* is an endomorphism which is a bijection. Now

Definition 2.1. *(Group representation) A representation of a group G on a complex vector space V is a homomorphism ρ*

$$\rho : G \to Aut(V) ,$$

where $Aut(V)$ is a family of automorphisms of V.

In other words, a representation of G on V is a map ρ such that for all $g \in G$ the map $\rho(g) : V \to V$ is linear and invertible and such that for all $g, g' \in G$ we have $\rho(gg')(v) = \rho(g)\rho(g')(v)$, $v \in V$. We say that V is a *G-module*. The dimension of V is called the *dimension of the representation*. It is customary (if slightly inaccurate) to call the vector space V itself a representation of G and write gv as a shorthand for $\rho(g)(v)$. If V is endowed with an inner product $\langle \cdot, \cdot \rangle$ then it is a *unitary representation* if the group acts on members of V unitarily, i.e. for any $g \in G$ and $v, u \in V$

$$\langle gv, gu \rangle = \langle v, u \rangle .$$

Next we define the *direct sum* and the *tensor product* of two vector spaces V and W. The direct sum $V \oplus W$ is a Cartesian product $V \times W$ with element-wise addition $(v_1, w_1) + (v_2, w_2) = (v_1 + v_2, w_1 + w_2)$ for $v_1, v_2 \in V$, $w_1, w_2 \in W$. If V and W are inner product spaces, so is its direct sum via $\langle (v_1, w_1), (v_2, w_2) \rangle = \langle v_1, v_2 \rangle + \langle w_1, w_2 \rangle$. Of course we have $\dim(V \oplus W) = \dim(V) + \dim(W)$.

Suppose now V and W are finite-dimensional; to construct a direct product $V \otimes W$ we associate to each pair (u, v) an element $u \otimes v$ via a bi-linear map $\phi : V \times W \to V \otimes W$. The map ϕ is unique up to an isomorphism. Consider a space of all linear combinations $F := \{\sum_n \alpha_n u_n \otimes v_n, (u_n, v_n) \in U \times V, \alpha_n \in \mathbb{C}\}$ and a subspace S generated by all the elements of F of the form

$$u \otimes w + v \otimes w - (u + v) \otimes w, \quad u \otimes w + u \otimes v - u \otimes (w + v) ,$$
$$(\alpha u) \otimes v - \alpha(u \otimes v), \quad u \otimes (\alpha v) - \alpha(u \otimes v)$$

for $\alpha \in \mathbb{C}$ and u, v, w in the appropriate spaces. Then the tensor product space is defined by taking $\phi(u,v) = (u, v + S)$, $(u,v) \in U \times V$, i.e. $U \otimes V$ is the quotient space $F \backslash S$. We can endow the new vector space $V \otimes W$ with a unique inner product such that, if $\{e_i\}_i$ and $\{f_j\}_j$ are orthonormal bases of V and W respectively, then $\{e_i \otimes f_j\}_{i,j}$ is an orthonormal basis of $V \otimes W$. This yields an alternative, somewhat more intuitive construction of $U \otimes V$, which, however, has a disadvantage of not being basis-free. It allows us to conclude that $\dim(V \otimes W) = \dim(V)\dim(W)$. If, in addition, vector spaces U and V are algebras, then the tensor product $U \otimes V$ can be given the structure of an algebra by defining $(u_1 \otimes v_1)(u_2 \otimes v_2) = (u_1 u_2) \otimes (v_1 v_2)$, for all $u_1, u_2 \in U$ and $v_1, v_2 \in V$.

The above two constructions generalise in a natural manner to more than two vector spaces. We write $\bigotimes^n V$ for the n-fold tensor product of a vector space V.

Suppose V and W are two finite-dimensional representations of G. Then the direct sum $V \oplus W$ and the direct product $V \otimes W$ are also representations of G; the former via

$$g(v \oplus w) = gv \oplus gw \ ,$$

and the latter via

$$g(v \otimes w) = gv \otimes gw \ .$$

We say that a subspace W of a representation V is an *invariant subspace* if it is invariant under the group action, i.e. $gw \in W$ for every $w \in W$ and $g \in G$. If V doesn't have any non-trivial invariant subspaces then we call V an *irreducible representation* (or an *irrep*), otherwise it is *reducible*. If a representation V can be written as a direct sum of invariant subspaces, then we say V is *completely reducible* or *decomposable*.

Now imagine two finite-dimensional irreps V and W of G, and suppose that the tensor product representation of V and W is completely reducible. Then one can write $V \otimes W$ as a direct sum

$$V \otimes W = \bigoplus_r V^r \ ,$$

where each V^r is an irreducible representation of G (note that the irreps on the RHS do not necessarily have multiplicities one). Since for each r V^r is a subspace of $V \otimes W$, the orthonormal basis vectors $\{v_k^r\}_k$ of V^r can be written as linear combinations of the basis vectors of $V \otimes W$

$$v_k^r = \sum_{i,j} \alpha_{i,j}^{r,k} e_i \otimes f_j \ .$$

The coefficients $\alpha_{i,j}^{r,k}$ are called *Clebsch-Gordan* or *Wigner coefficients*, as we will refer to them.

We are now ready to introduce *Lie groups*. A real (resp. complex) Lie group is a group with the compatible structure of a differentiable real (resp. complex) manifold and a group, i.e. the multiplication and the inverse group operations

$$\times : G \times G \to G \ ,$$
$$i : G \to G$$

are differentiable maps. Consider a group of all invertible real $n \times n$ matrices $GL_{\mathbb{R}}(n)$, called *the general linear group*. Each $n \times n$ matrix can be associated to a point in $\mathbb{R}^{n^2}$, matrix multiplication and matrix inversion operations are differentiable maps and so $GL_{\mathbb{R}}(n)$ is a real Lie group. The group of $n \times n$ complex invertible matrices $GL_{\mathbb{C}}(n)$ is a complex Lie group. So is the subgroup $SL_{\mathbb{C}}(n) = \{A \in GL_{\mathbb{C}}(n) : det(A) = 1\}$ called *the special linear group*. *The Unitary group* $U(n) = \{A \in GL_{\mathbb{C}}(n) : AA^* = A^*A = \mathbb{I}\}$ and *the special Unitary group* $SU(n) = \{A \in U(n) : det(A) = 1\}$ are real Lie groups. Here $*$ denotes conjugate transpose and $\mathbb{I}$ is an identity matrix of appropriate size. The above-mentioned groups are all *matrix groups*.

A fundamental role in the representation theory of Lie groups is played by Lie algebras. In general

Definition 2.2. *(Lie algebra) A real (resp. complex) Lie algebra $\mathfrak{g}$ is a real (resp. complex) vector space together with a bi-linear anti-symmetric operator $[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, called the Lie bracket, satisfying the Jacobi identity*

$$[x,[y,z]] + [y,[z,x]] + [z,[x,y]] = 0 \ .$$

Every linear Lie group has a real Lie algebra associated to it (see e.g [22, Thm. 10.5.III]). The purpose of a real Lie algebra is to capture and describe 'local' properties of the associated Lie group. More specifically a *real* Lie algebra of a Lie group is a tangent space of the group at the identity. Such tangent space naturally has the structure of a Lie algebra (see, for example, [35, §8.1]). Informally speaking, close to the identity every member of the group G can be approximated by a member of the associated real Lie algebra $\mathfrak{g}$. Lie algebra of a Lie group doesn't however capture the global information about the group; there exist groups with isomorphic Lie algebras which are not themselves isomorphic.

Note that matrices forming a real Lie algebra of a matrix group need not themselves have real entries. One can easily *complexify* a Lie algebra to obtain the corresponding complex Lie algebra.

A homomorphism of two Lie algebras $\mathfrak{g}$ and $\mathfrak{g}'$ is a linear map $\rho : \mathfrak{g} \to \mathfrak{g}'$ which preserves the Lie bracket, i.e.

$$\rho([X,Y]) = [\rho(X), \rho(Y)], \quad X, Y \in \mathfrak{g} .$$

Definition 2.3. *(Representation of a Lie algebra) A representation of a Lie algebra on a finite-dimensional vector space V is a homomorphism ρ*

$$\rho : \mathfrak{g} \to End(V) ,$$

where $End(V)$ stands for the collection of all endomorphisms of V.

We say that a representation V of an algebra $\mathfrak{g}$ is a module with respect to that algebra. Every representation of a group induces a representation of the corresponding Lie algebra. The converse however is not as straightforward, but is true in some situations. To see how we can relate a representation of a Lie algebra to a representation of a Lie group we need to introduce the *exponential map*. Let $\{g(t) \in G : t \in \mathbb{R}\}$ be a subgroup of a linear Lie group G such that $g(t)g(s) = g(t+s)$ for all $t,s \in \mathbb{R}$. We call this a *one-parameter subgroup of G*. Now

Theorem 2.4. *([22, Thm. 10.5.V]) Every element X of a real Lie algebra $\mathfrak{g}$ of a linear Lie group G can be associated with a one-parameter subgroup of G defined by*

$$g(t) = \exp(tX) .$$

The map $\exp : \mathfrak{g} \to G$ *is called the exponential map.*

For compact semisimple Lie groups the exponential map is onto [74, Cor. VIII.1.2]. Notably $SU(n)$ is an example of a compact, semisimple Lie group [74, Ch. VIII.4, p. 174]; we will use this fact later in the chapter. Finally we note that for the matrix groups the exponential map is the usual matrix exponentiation.

Recall that if V and U are finite-dimensional representations of a Lie group G so is the direct product $V \otimes U$. The representations of the group on U and V induce representations of the Lie algebra on U and V. We would like an action of the group on $U \otimes V$ to induce a representation of the algebra on $U \otimes V$. This is done by defining the action of the algebra on the elements of the space $V \otimes U$ by

$$X(v \otimes u) = (Xv) \otimes u + v \otimes (Xu)$$

for $v \in V$, $u \in U$ and $X \in \mathfrak{g}$. Again if a tensor product of two representations is completely reducible, it can be expressed as a direct product of irreps, i.e. we can write $V \otimes W \simeq \bigotimes_r V^r$. Note also that the Wigner coefficients of a Lie algebra describing the embedding $V^r \subset V \otimes W$ (which are defined in an analogous manner to the Wigner coefficients of a Lie group) are the same as the Wigner coefficients of the corresponding Lie group describing the same embedding [22, pp. 412-413].

Finally we introduce the universal enveloping algebra $U(\mathfrak{g})$ of a Lie algebra $\mathfrak{g}$. We will see later, when we introduce some basic Quantum Probability theory, that many observables of interest are members of appropriate universal enveloping algebras. Let $\mathfrak{g}$ be an algebra over complex numbers and define $T^0(\mathfrak{g}) = \mathbb{C}$, $T^1(\mathfrak{g}) = \mathfrak{g}$ and

$T^n(\mathfrak{g}) = \bigotimes^n \mathfrak{g}$ for $n \geq 1$. Let $T(\mathfrak{g})$ be the vector space defined by

$$T(\mathfrak{g}) = \bigoplus_k T^k(\mathfrak{g}) = \mathbb{C} \oplus \mathfrak{g} \oplus (\mathfrak{g} \otimes \mathfrak{g}) \oplus (\mathfrak{g} \otimes \mathfrak{g} \otimes \mathfrak{g}) \otimes \ldots$$

with multiplication law given by the tensor product, i.e.

$$(x_1 \otimes \ldots \otimes x_n)(x_{n+1} \otimes \ldots \otimes x_m) = x_1 \otimes \ldots \otimes x_n \otimes x_{n+1} \otimes \ldots \otimes x_m$$

for $x_i \in \mathfrak{g}$, $1 \leq i \leq m$. We call $T(\mathfrak{g})$ the *tensor algebra* of $\mathfrak{g}$. Let $I(\mathfrak{g})$ be the two-sided ideal generated by all the elements of the form

$$[x, y] - x \otimes y + y \otimes x, \qquad x, y \in \mathfrak{g} ,$$

i.e. $I(\mathfrak{g})$ is a subspace of $T(\mathfrak{g})$ with $x - y \in I(\mathfrak{g})$ if $x, y \in I(\mathfrak{g})$ and $zx, xz \in I(\mathfrak{g})$ for all $x \in I(\mathfrak{g})$ and $z \in T(\mathfrak{g})$. Then the *universal enveloping algebra of* $\mathfrak{g}$, denoted $U(\mathfrak{g})$, is the associative algebra given by the quotient space of the tensor algebra $T(\mathfrak{g})$ by the two-sided ideal $I(\mathfrak{g})$, i.e. $U(\mathfrak{g}) = T(\mathfrak{g}) \backslash I(\mathfrak{g})$. Evidently, $\mathfrak{g} \subset U(\mathfrak{g})$.

Definition 2.5. *(Representation of universal enveloping algebra) Let $U(\mathfrak{g})$ be a universal enveloping algebra of a Lie algebra $\mathfrak{g}$. A representation of $U(\mathfrak{g})$ on a finite-dimensional complex vector space V is a homomorphism*

$$\widetilde{\rho} : U(\mathfrak{g}) \rightarrow End(V) .$$

Let ρ be a representation of the Lie algebra $\mathfrak{g}$. Then ρ extends uniquely to a representation of the universal enveloping algebra by

$$\widetilde{\rho}(x_1 \otimes \ldots \otimes x_n) = \rho(x_1) \otimes \ldots \otimes \rho(x_n) ,$$

for $x_i \in \mathfrak{g}$, $1 \leq i \leq n$. Conversely, from every representation $\widetilde{\rho}$ of $U(\mathfrak{g})$ we can obtain a representation ρ of $\mathfrak{g}$ by letting $\rho(x) = \widetilde{\rho}(x)$ for $x \in \mathfrak{g}$.

2.2 Representations of $\mathfrak{sl}(2)$ and $U(\mathfrak{sl}(2))$

We are now ready to discuss concrete examples. We start with $\mathfrak{sl}_{\mathbb{C}}(2)$, the (complex) Lie algebra of the special linear group $SL(2)$. To construct a certain discrete bi-variate Markov chain in Chapter 3, we need to study finite-dimensional irreducible representations of $\mathfrak{sl}_{\mathbb{C}}(2)$ and in particular the Wigner coefficients describing the decomposition of the tensor products of the representations into irreps.

The Lie algebra $\mathfrak{sl}_{\mathbb{C}}(2) := \mathfrak{sl}(2)$ is a complex vector space of 2×2 complex matrices with trace zero spanned by three generators

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad Y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},$$

which satisfy the commutation identities

$$[H,X] = 2X, \qquad [H,Y] = -2Y, \qquad [X,Y] = H\ , \tag{2.1}$$

where the *Lie bracket* $[\cdot,\cdot]$ is given by the commutator

$$[A,B] = AB - BA; \qquad A,B \in \mathfrak{g}$$

and the multiplication rule by the usual matrix multiplication.

It is well known that for each integer $n \geq 0$ there is a unique, up to an isomorphism, irreducible representation of $\mathfrak{sl}(2)$ of dimension $n+1$; we denote it by V_n. In particular $V_0 = \{1\}$ is a *trivial representation* with $\rho_{V_0}(a) = 1$ for all $a \in \mathfrak{sl}(2)$ and $V_1 := V \simeq \mathbb{C}^2$ is a *natural representation* with $\rho_{V_1}(a) = a$, $a \in \mathfrak{sl}(2)$.

Suppose V is any finite-dimensional representation of $\mathfrak{sl}(2)$. A scalar α is called a *weight* and a vector $v \in V$ the associated *weight vector* if $Hv = \alpha v$, i.e. if the action of H on V is diagonalisable with eigenvalue α and the corresponding eigenvector v. A vector $v \in V$ is called the *highest* (resp. *lowest*) *weight vector* if additionally $Xv = 0$ (resp. $Yv = 0$). Every finite-dimensional $\mathfrak{sl}(2)$-representation has the highest (and lowest) weight vector (see [48, Prop. V.4.2]).

In what follows we shall identify V_n as a subspace of the n-fold tensor product $\bigotimes^n V$ of the natural representation. Let $\{e_0, e_1\} := \{\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}\}$ be the canonical orthonormal basis of $\mathbb{C}^2 \simeq V$. Then e_0 is the lowest weight vector of V_1 with weight -1. Define e_k^n as the image of the action of the k^{th} power of X on $e_0^n := \otimes^n e_0$, i.e.

$$e_k^n := \sqrt{\frac{(n-k)!k!}{n!}} X^k(\otimes^n e_0) = \sqrt{\frac{(n-k)!k!}{n!}} \sum_\sigma e_{\sigma_1} \otimes \ldots \otimes e_{\sigma_n} ,$$

where the summation is over all distinct permutations of indices of e_i's such that the number of i's equal to 1 is k. The constant involving factorials in front of X^k ensures that the vector e_k^n has unit length. Then $Xe_k^n = 0$ for all $k \geq n$ and the set of vectors $(e_0^n, \ldots, e_n^n)$ is orthonormal. Moreover one calculates

$$\begin{aligned} He_k^n &= (2k-n)e_k^n , \\ Xe_k^n &= \sqrt{(n-k)(k+1)}e_{k+1}^n , \\ Ye_k^n &= \sqrt{k(n-k+1)}e_{k-1}^n . \end{aligned} \tag{2.2}$$

The invariant space spanned by the vectors $\{e_i^n; 0 \leq i \leq n\}$ is isomorphic to the irreducible representation V_n of dimension $n+1$ (see [48, Thm. V.4.4] or [35, Claim 11.4]). The action of H on V_n is diagonalisable with eigenvalues $(-n, -n+2, \ldots, n-2, n)$ with corresponding eigenvectors $(e_0^n, \ldots, e_n^n)$. In particular the highest weight vector e_n^n corresponds to the highest weight n. Thus constructed vector spaces V_n, $n \geq 0$, constitute *all* finite-dimensional irreducible representations of $\mathfrak{sl}_{\mathbb{C}}(2)$.

It is known that every finite-dimensional representation of $\mathfrak{sl}(2)$ is completely reducible (this follows from the fact that $\mathfrak{sl}(2)$ is semisimple and, for example, [35, Thm. 9.19]), that is, it can be written as a direct sum of irreps. Consider a tensor product of V_n with the natural representation V. What is the direct sum decomposition of $V_n \otimes V$? The product space has the orthonormal basis $\{e_k^n \otimes e_0, e_k^n \otimes e_1; 0 \leq k \leq n\}$, and so one easily finds that the action of H on the space is diagonalisable with eigenvalues $(-n+1, -n+3, \ldots, n-3, n-1)$ with multiplicity two and $(-n-1, n+1)$ with multiplicity one. Thus, $V_n \otimes V$ must contain a copy of V_{n-1} and a copy of V_{n+1} and we have

$$V_n \otimes V \simeq V_{n-1} \oplus V_{n+1} \; . \tag{2.3}$$

We would like to calculate Wigner coefficients describing embeddings $V_{n-1}, V_{n+1} \subset V_n \otimes V$.

Proposition 2.6. *Let* $\{f_p; 0 \leq p \leq n-1\}$*, resp.* $\{u_k; 0 \leq k \leq n+1\}$*, be an orthonormal basis of* V_{n-1}*, resp.* V_{n+1}*, in the decomposition* (2.3)*, such that* $Hf_p = (2p-(n-1))f_p$ *and* $Hu_k = (2k-(n+1))u_k$*. Then*

$$\begin{aligned} f_p &= \sqrt{\frac{n-p}{n+1}} e_p^n \otimes e_1 - \sqrt{\frac{p+1}{n+1}} e_{p+1}^n \otimes e_0 \; , \\ u_k &= \sqrt{\frac{k}{n+1}} e_{k-1}^n \otimes e_1 + \sqrt{\frac{n-k+1}{n+1}} e_k^n \otimes e_0 \end{aligned} \tag{2.4}$$

for $0 \leq p \leq n-1$ *and* $0 \leq k \leq n+1$*.*

Proof. By construction f_k and u_k are weight vectors with weights $2k-(n-1)$ and $2k-(n+1)$ respectively. But for $1 \leq k \leq n$

$$H(e_k^n \otimes e_0) = (2k-n-1) e_k^n \otimes e_0 \quad \text{and} \quad H(e_k^n \otimes e_1) = (2k-n+1) e_k^n \otimes e_1 \; ,$$

and so we must have

$$\begin{gathered} u_{n+1} = e_n^n \otimes e_1, \qquad u_0 = e_0^n \otimes e_1 \; , \\ u_k = \alpha e_{k-1}^n \otimes e_1 + \beta e_k^n \otimes e_0 \quad \text{and} \quad f_{k-1} = \alpha' e_{k-1}^n \otimes e_1 + \beta' e_k^n \otimes e_0 \; , \end{gathered}$$

for some $\alpha, \beta, \alpha', \beta' \in \mathbb{C}$. To calculate the constants $\alpha, \beta, \alpha', \beta'$ recall that in order to span V_{n-1} and V_{n+1} the sets of vectors $(f_p; 0 \leq p \leq n-1)$ and $(u_k; 0 \leq k \leq n+1)$ must satisfy relations (2.2). In particular, the action of X must take f_k (resp. u_k) to f_{k+1} (resp. u_{k+1}) and kill f_{n-1} (resp. u_{n+1}). Thus successively applying X to f_0 (resp. u_0) we find that this happens precisely when the constants $\alpha, \alpha', \beta, \beta'$ are as in (2.4). Note also that one can achieve the same result by applying powers of Y to f_{n-1} and u_{n+1} instead. $\square$

The universal enveloping algebra $U(\mathfrak{sl}(2))$ is isomorphic to a complex algebra generated by the three abstract elements H, X and Y satisfying relations (2.1). Hence the above analysis is still valid and shows that finite-dimensional irreducible $U(\mathfrak{sl}(2))$-modules are isomorphic to representations V_n of $\mathfrak{sl}(2)$ described above (see [48, Ch. V.3]).

2.3 Quantum enveloping algebra $U_q(\mathfrak{sl}(2))$

It is possible to construct a one-parameter deformation, or *quantisation*, of $U(\mathfrak{sl}(2))$. It is called the *quantum enveloping algebra* and denoted by $U_q := U_q(\mathfrak{sl}(2))$ (for more on quantisation of algebras see, for example, [21, Ch. 9.1]). Parameter q is complex and must not be a root of unity; in our case $q \in (0,1)$. The quantum enveloping algebra $U_q(\mathfrak{sl}(2))$ is a complex algebra with four generators X, Y, q^H and q^{-H} satisfying the following identities

$$\begin{aligned} q^H q^{-H} =& q^{-H} q^H = 1 \ , \\ q^H X q^{-H} = q^2 X, \quad & q^H Y q^{-H} = q^{-2} Y \ , \\ [X, Y] =& \frac{q^H - q^{-H}}{q - q^{-1}} \ . \end{aligned} \tag{2.5}$$

Letting q tend to 1 in the above, one can recover identities (2.1) satisfied by the generators X, Y and H of $U(\mathfrak{sl}(2))$.

Coproduct, or *comultiplication*, is a linear function $\triangle : U_q \to U_q \otimes U_q$ (which, together with the counit and certain relations between the two maps, endows U_q with the coalgebra structure) is defined by the following relations

$$\begin{aligned} \triangle(q^{\pm H}) = q^{\pm H} \otimes q^{\pm H} \ , \\ \triangle(Z) = Z \otimes q^{-H/2} + q^{H/2} \otimes Z; \quad Z \in \{X, Y\} \ . \end{aligned}$$

Note. This definition of $U_q(\mathfrak{sl}(2))$ is an $n = 2$ adaptation of the definition of $U_q(\mathfrak{gl}(n)$ of Jimbo, Date and Miwa [24] and so the coproduct action differs from a more conventional definition $\triangle(Z) = Z \otimes 1 + q^H \otimes Z$, $Z \in \{X, Y\}$, found in most books.

Coproduct defines action of U_q on tensor products of representations: if ρ_1 and ρ_2 are two representations of U_q on vector spaces V_1 and V_2 respectively, then representation on the direct product $V_1 \otimes V_2$ is defined by

$$(\rho_1 \otimes \rho_2)\triangle : U_q \to End(V_1 \otimes V_2)\ .$$

We now describe all finite-dimensional irreducible representations, or modules, of U_q. For each integer $n \geq 0$ there exist two irreducible representations of U_q of dimension $n+1$; we denote them by $V_{n,+}$ and $V_{n,-}$. The highest (resp. lowest) weight vector of a finite-dimensional module V is a non-zero vector $v \in V$ such that $q^H v = \alpha v$ and $Xv = 0$ (resp. $Yv = 0$). Every non-zero finite-dimensional U_q-module has the highest weight vector (see [48, Prop. VI.3.3]). Let $V_{n,\pm}$ be a vector space with orthonormal basis $\{e_0^{n,\pm}, ..., e_n^{n,\pm}\}$ and define the action of the generators as follows

$$\begin{aligned} q^H e_k^{n,\pm} &= \pm q^{2k-n} e_k^{n,\pm}\ , \\ X e_k^{n,\pm} &= \sqrt{[n-k]_q[k+1]_q} e_{k+1}^{n,\pm}\ , \\ Y e_k^{n,\pm} &= \pm\sqrt{[k]_q[n-k+1]_q} e_{k-1}^{n,\pm}\ , \end{aligned} \tag{2.6}$$

where for $k > 0$

$$[k]_q = \frac{q^k - q^{-k}}{q - q^{-1}}\ .$$

In particular $e^{n,\pm}$ is the highest weight vector with weight $\pm q^n$. One can check that thus defined actions on $V_{n,\pm}$ agree with (2.5). Vector spaces $V_{n,\pm}$ are unique $(n+1)$-dimensional representations of U_q up to an isomorphism ([48, Thm. V.4.4]).

From now on we are only interested in representations $V_{n,+}$. The branching rule for $V_{n,+} \otimes V$ is the same as in the classical case (2.3), i.e.

$$V_{n,+} \otimes V \simeq V_{n-1,+} \otimes V_{n+1,+} \ .$$

Again to see this, it is enough to note that the action of q^H on the tensor space $V_{n,+} \otimes V$ is diagonalisable with eigenvalues $(q^{2k-n-1}; 1 \leq k \leq n)$ with multiplicity two and (q^{-n-1}, q^{n+1}) with multiplicity one.

Proposition 2.7. *Let $\{f_p; 0 \leq p \leq n-1\}$, resp. $\{u_k; 0 \leq k \leq n+1\}$, be an orthonormal basis of $V_{n-1,+}$, resp. $V_{n+1,-}$, in the decomposition of $V_{n,+} \otimes V$ above, such that $Hf_p = q^{2p-(n-1)} f_p$ and $Hu_k = q^{2k-(n+1)} u_k$. Then*

$$\begin{aligned} f_p &= q^{(p+1)/2} \sqrt{\frac{[n-p]_q}{[n+1]_q}} e^n_p \otimes e_1 - q^{(p-n)/2} \sqrt{\frac{[p+1]_q}{[n+1]_q}} e^n_{p+1} \otimes e_0 \ , \\ u_k &= q^{(k-n-1)/2} \sqrt{\frac{[k]_q}{[n+1]_q}} e^n_{k-1} \otimes e_1 + q^{k/2} \sqrt{\frac{[n-k+1]_q}{[n+1]_q}} e^n_k \otimes e_0 \ , \end{aligned} \tag{2.7}$$

for $0 \leq p \leq n-1$ and $0 \leq k \leq n+1$.

Proof. The proof proceeds exactly like the proof of Proposition 2.6. □

Note that by letting q tend to 1 we can recover expressions (2.4).

2.4 Some combinatorics

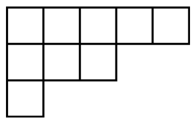
Before we talk about the enveloping algebra $U(\mathfrak{gl}(n))$, we state some relevant definitions and results from combinatorics. For proofs of results in this section and more details see books by Stanley [75] or Sagan [73].

A *partition* is a sequence of positive numbers $(\lambda_1, \lambda_2, \lambda_3 ...)$ with $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq ...$ and a number $n \geq 1$, called the *size of the partition*, such that $\lambda_m = 0$ exactly for all $m > n$. It is convenient to write $\lambda = (\lambda_1, ..., \lambda_n)$ to designate such partition. We denote by $\mathscr{P}_n$ the collection of all partitions of size n. If $|\lambda| = \sum_i \lambda_i = a$ we write $\lambda \vdash a$. For two partitions λ and μ we write $\mu \prec \lambda$ if

$$\lambda_1 \geq \mu_1 \geq \lambda_2 \geq \mu_2 \geq \dots$$

and $\mu \nearrow \lambda$ to say that $\mu \prec \lambda$ and $|\lambda| = |\mu| + 1$. If $\mu \prec \lambda$, we say that μ *is interlaced with* λ.

A *Young diagram* is a left-aligned array of boxes with weakly increasing row lengths. Each diagram is characterised by its *shape*, a vector of the row lengths. The number of boxes in a diagram is called its *size*. Evidently if λ is a shape of a Young diagram of size a, then $\lambda \vdash a$. Figure below, for example, depicts a Young diagram of size 9 and with shape given by a vector $\lambda = (5, 3, 1) \vdash 9$.



A *filling of a Young diagram* with integers from $[n] := \{1, 2, \dots, n\}$ such that entries are weakly increasing along the rows and strictly increasing down the columns is called a *semi-standard Young tableau* or simply a *tableau*. A semi-standard Young tableau which entries are all distinct is called a *standard Young tableau*. A tableau with entries in $[n]$ has at most n rows. The *type* of a Young tableau with values in $[n]$ is a vector $(x_1, \dots, x_n)$ where x_i is the number of i's in the tableau. Below is a semi-standard filling of the Young diagram above with numbers from $[4]$; the type of the tableau is $(2, 2, 2, 3)$.

1	1	2	3	4
2	3	4		
4				

We call a *word* a sequence of positive integers $\{a_1, a_2, \dots, a_k\}$ such that $a_i \in [n]$ for $1 \leq i \leq k$. With any word $a \in [n]^k$ one can associate a pair of Young tableaux (P, Q) of size k and entries in $[n]$. The P-tableau is semistandard and is constructed via a so called *row insertion* or *row bumping* algorithm. The algorithm proceeds as follows.

- At step one add a box containing number a_1.

- At time step $k > 1$ consider a_k. If the rightmost box in the first row has a number smaller than or equal to a_k, append a_k at the end. Otherwise insert a_k instead of the rightmost number bigger than a_k and place the displaced number in the second row via the same procedure. Proceed until no boxes can be displaced.

The Q-tableau, called a *recording tableau*, is a standard Young tableau of the same shape as the P-tableau and is constructed by placing in each box the number of the step at which this box was added to the P-tableau. The *Robinson-Schensted correspondence*, or the *RS correspondence*, states that words in $[n]^k$ and pairs of Young tableaux (P,Q), where P is semistandard tableau of size k with entries from $[n]$ and Q is a standard tableau of the same shape, are in *one-to-one correspondence*. In the next chapter we will look at the Robinson-Schensted algorithm with random input.

One can also apply the Robinson-Schensted algorithm with *column insertion* instead of the row-insertion; it is defined analogously to the row insertion with the role of rows and columns interchanged. The resulting correspondence between random words and pairs (P,Q) is also a bijection.

With each semistandard Young tableau one can associate a so-called *Gelfand-Cetlin pattern* (sometimes spelled Gelfand-Tsetlin pattern), abbreviated GC-pattern.

Definition 2.8. *(Gelfand-Cetlin pattern) A Gelfand-Cetlin pattern of depth n is an array of positive real numbers $m = (m^n, ..., m^1) \in \mathbb{R}_+^n \times ... \times \mathbb{R}_+$ such that $m^1 \succ m^2 \succ ... \succ m^n$, i.e. such that the following interlacing inequalities*

$$m_j^i \geq m_j^{i+1} \geq m_{j+1}^i \tag{2.8}$$

hold. Schematically a GC-pattern can be expressed as a triangular array

$$\begin{array}{ccccccccc}
m_1^n & & m_2^n & \cdots & m_{n-1}^n & & m_n^n \\
& m_1^{n-1} & & \cdots\cdots\cdots\cdots & & m_{n-1}^{n-1} & \\
& \ddots & & \cdots\cdots\cdots\cdots & & \iddots & \\
& & m_1^2 & & m_2^2 & & \\
& & & m_1^1 & & &
\end{array} .$$

In what follows we will primarily consider GC-patterns with natural numbers as entries. One can construct a GC-pattern of depth n from a semi-standard Young tableau with entries in $[n]$ as follows. The top row of the pattern is taken to be the shape of the tableau. If we remove all the boxes filled with number n, by construction of the tableau, we obtain a semi-standard Young tableau again. Moreover, the shape of the new tableau will be interlaced with the shape of the original one. Take the shape of the new tableau to be the second row of the pattern. We proceed by stripping the Young tableau of boxes filled with consecutive numbers, at each step taking the shape of the new tableau to be the next row of the pattern. Interlacing conditions (2.8) are satisfied automatically because of the definition of a Young tableau. We see that there is a one-to-one correspondence between Geldand-Cetlin patterns of depth n and semi-standard Young tableaux with entries in $[n]$, or equivalently the nested sequences formed by their rows. For example

$$\begin{array}{ccccccc} 5 & & 3 & & 1 & & 0 \\ & 4 & & 2 & & 0 & \\ & & 3 & & 1 & & \\ & & & 2 & & & \end{array}$$

is the GC-pattern associated with the Young tableau above.

We will denote by $\mathscr{K}_\lambda$ the collection of all GC-patterns with top row $\lambda = (\lambda_1, ..., \lambda_n)$, and call $\mathscr{K}_\lambda$ the *Gelfand-Cetlin cone*. We will need the following classical result

Theorem 2.9. *(see, for example, [35, Thm. 6.3]) The number of Gelfand-Cetlin patterns with top row $\lambda = (\lambda_1, ..., \lambda_n)$ is*

$$|\mathscr{K}_\lambda| = \prod_{i<j} \frac{\lambda_i - \lambda_j - i + j}{j - i} \ . \tag{2.9}$$

For a partition λ the *Vandermonde determinant* $\triangle(\lambda)$ is given by

$$\triangle(\lambda) = \det(\lambda_i^{k-j}) = \prod_{i<j} (\lambda_i - \lambda_j) \ . \tag{2.10}$$

Hence, if we let $\widetilde{\lambda}_i = \lambda_i - i$, then we can rewrite (2.9) as

$$|\mathscr{K}_\lambda| = \frac{1}{Z_n} \triangle(\widetilde{\lambda}) \ , \tag{2.11}$$

where $Z_n := 1/\prod_{i<j}(i-j)$.

2.5 Representation theory of universal enveloping algebra $U(\mathfrak{gl}(\mathfrak{n}))$ and quantum algebra $U_q(\mathfrak{gl}(n))$

In chapter 5 we will construct an n-dimensional generalisation of the family of discrete bivariate chains considered in the next chapter. To do that we need to study representation theory of $U(\mathfrak{gl}(n))$, the universal enveloping algebra of the Lie algebra $\mathfrak{gl}_{\mathbb{C}}(n) := \mathfrak{gl}(n)$ associated to the general linear group of degree n, and its quantum counterpart $U_q(\mathfrak{gl}(n))$. The Lie algebra $U(\mathfrak{gl}(n))$ is a complex associative unital algebra spanned by the generators $\{E_{ij}; 1 \leq k \leq n\}$ satisfying the commutation relation

$$[E_{ij}, E_{kl}] = \delta_{jk}E_{il} - \delta_{il}E_{kj} \ . \tag{2.12}$$

For each partition λ with at most n parts there exists an irreducible representation V_λ of $U(\mathfrak{gl}(n))$ of dimension $|\mathscr{K}_\lambda|$. The space V_λ is spanned by vectors parameterised by the fillings of a Young tableau of shape λ or, alternatively, by all the GC-patterns with top row λ. We call this basis the *Gelfand-Cetlin basis*, and equip V_λ with an inner product with respect to which it is orthonormal. Write e_m for a vector from the GC basis corresponding to the pattern m. It suffices to define the actions of $E_{k,k+1}, E_{k+1,k}$ and E_{kk} on V_λ for all k. Actions on V_λ of all the other generators can be deduced from the actions of $E_{kk}, E_{k,k+1}, E_{k+1,k}$ and the commutator relations (2.12). Indeed, we have

$$[E_{k,k+1}, E_{k+1,k+h}] = E_{k,k+h}, \quad [E_{k+h,k+1}, E_{k+1,k}] = E_{k+h,k} \ . \tag{2.13}$$

for all $h \geq 1$.

For a pattern $m \in \mathscr{K}_\lambda$ we define

$$E_{kk} e_m = \left(\sum_{k=1}^{i} m_k^i - \sum_{k=1}^{i-1} m_k^{i-1} \right) e_m \ , \tag{2.14a}$$

$$E_{k,k+1} e_m = \sum_{\widehat{m}}^{(k)} c_k(\widehat{m}, m) e_{\widehat{m}}, \quad E_{k+1,k} e_m = \sum_{\widehat{m}}^{(k)} c_k(m, \widehat{m}) e_{\widehat{m}} \ , \tag{2.14b}$$

where $\sum_{\widehat{m}}^{(k)}$ stands for a summation over all $\widehat{m}$ such that $\widehat{m}^i = m^i$ if $i \neq k$. Let $l_j^i = m_j^i - j$. The coefficient $c_k(m, \widehat{m})$, for $m, \widehat{m} \in \mathscr{K}_\lambda$, is defined by

$$c_k(m, \widehat{m}) = \left(- \frac{\prod_{i=1}^{k-1} (l_i^{k-1} - l_j^k) \prod_{i=1}^{k+1} (l_i^{k+1} - l_j^k + 1)}{\prod_{i \neq j} (l_i^k - l_j^k)(l_i^k - l_j^k + 1)} \right)^{\frac{1}{2}}$$

and is non-zero only if $\widehat{m}_j^k = m_j^k - 1$ for some j and $\widehat{m}_p^i = m_p^i$ for all $(i,p) \neq (k,j)$, i.e. the pattern $\widehat{m}$ is obtained from the pattern m by subtracting 1 from a single element m_j^k. We see that every vector $e_m \in V_\lambda$, $m \in \mathscr{K}_\lambda$, is a weight vector with the associated weight $\sum_k m_k^i - \sum_k m_k^{i-1}$. In particular, the vector e associated to the Young tableau with only k's in the k^{th} row, or equivalently to the GC-pattern

$$\begin{array}{ccccccccc}
\lambda_1 & & \lambda_2 & \dots & \lambda_{n-1} & & \lambda_n \\
& \lambda_1 & & \dots\dots\dots\dots\dots & & \lambda_{n-1} & \\
& \ddots & & \dots\dots\dots\dots\dots & & \iddots & \\
& & \lambda_1 & & \lambda_2 & & \\
& & & \lambda_1 & & &
\end{array}$$

is the highest weight vector with $E_{ii} e = \lambda_i e$ and $E_{ij} e = 0$ for all $i < j$.

The algebra $U(\mathfrak{gl}(n-1))$ is naturally embedded in $U(\mathfrak{gl}(n))$ via mappings $E_{ij} \to E_{ij}$ for $1 \leq i, j \leq n-1$. Restricted to the action of $U(\mathfrak{gl}(n-1))$, V_λ contains all the $U(\mathfrak{g}(n-1))$ irreducibles parameterised by $\mu = (\mu_1, ..., \mu_{n-1})$ such that $\mu \prec \lambda$ ([5, Ch. 8.8, Thm. 1]). We denote these invariant subspaces by V_λ^μ to stress the fact that they are embedded in V_λ. We write

$$V_\lambda \simeq \bigoplus_{\mu \prec \lambda} V_\lambda^\mu \ . \tag{2.15}$$

Each irreducible module on the right-hand side appears exactly once.

As before we are particularly interested in the tensor product representation $V_\lambda \otimes V$, where $V := V_{(1,0,0\dots)} = \mathbb{C}^n$ is the natural representation of $U(\mathfrak{gl}(n))$, and its decomposition into the direct sum of irreducibles. The branching rule is given as follows

$$V_\lambda \otimes V \simeq \bigoplus_{\lambda':\lambda\nearrow\lambda'} V_{\lambda'} \,. \tag{2.16}$$

Again the embedding of each $V_{\lambda'}$ in the space $V_\lambda \otimes V$ is described by the Wigner coefficients, although the formula is much more complicated than in the case $n = 2$ described in Section 2.2. Here we are using the notation and Wigner coefficients given in [24] for the case when $q = 1$. Suppose λ' is obtained from λ by adding a box at row p and let $m \in \mathscr{K}_{\lambda'}$. Then for $(i_n, \dots, i_j)$ with $i_n = p$, $1 \le i_k \le k$ and $1 \le j \le k \le n$, define a GC-pattern $\widehat{m} := (m; i_n, \dots, i_j) \in \mathscr{K}_\lambda$ such that $\widehat{m}^k_i = m^k_i$ for all $i \neq i_k$ and $\widehat{m}^k_{i_k} = m^k_{i_k} - 1$, $j \le k \le n$. In other words m is the GC-pattern associated with the Young tableau obtained by adding a box with number j at row p to the Young tableau associated with pattern $(m; i_n, \dots, i_j)$. When a box filled with number j is added, entries on at most top $n - j + 1$ rows change with exactly one entry increasing by 1 in each row. Each index i_k indicates which entry exactly changes in row k. For any $e'_m \in V_{\lambda'}$, $\lambda \nearrow \lambda'$, the Clebsch-Gordan formula reads

$$e'_m = \sum_{j=1}^{n} \sum_{(i_n,\dots,i_j)} w((m; i_n, \dots, i_j)) e_{(m;i_n,\dots,i_j)} \otimes e_j \,, \tag{2.17}$$

where $e_{(m;i_n,\dots,i_j)} \in V_\lambda$ and $\{e_1, \dots, e_n\}$ is the canonical basis of $\mathbb{C}^n$. The Wigner coefficients $w((m; i_n, \dots, i_1)) \equiv \langle e'_m, e_{(m;i_n,\dots,i_j)} \otimes e_j \rangle$ are given by

$$\begin{aligned} &w((m; i_n, \dots, i_j)) \\ &= w^{(1)} \left(\begin{array}{ccc|c} m^j_1 & \cdots & m^j_j & i_j \\ m^{j-1}_1 & \cdots & m^{j-1}_{j-1} & \end{array} \right) \prod_{k=j+1}^{n} w^{(2)} \left(\begin{array}{ccc|c} m^k_1 & \cdots & m^k_k & i_k \\ m^{k-1}_1 & \cdots & m^{k-1}_{k-1} & i_{k-1} \end{array} \right) \,, \end{aligned} \tag{2.18}$$

where $w^{(1)}$ and $w^{(2)}$ are the *reduced Wigner coefficients* defined by

$$\begin{aligned} w^{(1)}\left(\begin{array}{ccc|c} x_1 & \cdots & x_n & i \\ y_1 & \cdots & y_{n-1} & \end{array}\right) &= \left(\frac{\prod_{k\le n-1}(\widetilde{y}_k-\widetilde{x}_i)}{\prod_{k\le n,k\ne i}(\widetilde{x}_k-\widetilde{x}_i+1)}\right)^{\frac{1}{2}} , \\ w^{(2)}\left(\begin{array}{ccc|c} x_1 & \cdots & x_n & i \\ y_1 & \cdots & y_{n-1} & j \end{array}\right) &= \left(\prod_{k\le n,k\ne i}\frac{\widetilde{x}_k-\widetilde{y}_j+1}{\widetilde{x}_k-\widetilde{x}_i+1}\prod_{k\le n-1,k\ne j}\frac{\widetilde{y}_k-\widetilde{x}_i}{\widetilde{y}_k-\widetilde{y}_j}\right)^{\frac{1}{2}} , \end{aligned} \tag{2.19}$$

where $\widetilde{x}_k = x_k - k$ and $\widetilde{y}_k = y_k - k$.

To finish the discussion of representations of $U(\mathfrak{gl}(n))$ we prove

Lemma 2.10. *For any $\lambda = (\lambda_1, ..., \lambda_n)$ a $U(\mathfrak{gl}(n))$-module V_λ is a representation of $SU(n)$, the special Unitary group of degree n. Moreover, the action of the group on V_λ is unitary with respect to the inner product on V_λ which makes the Gelfand-Cetlin basis orthonormal, i.e.*

$$\langle g e_m, g e_{\widehat{m}} \rangle = \langle e_m, e_{\widehat{m}} \rangle$$

for all $e_m, e_{\widehat{m}} \in V_\lambda$ and $g \in SU(n)$.

which we will need in Chapter 5.

Proof. We start by recalling that any representation of the universal enveloping algebra can be restricted to a representation of the corresponding Lie algebra. Hence for any partition λ of at most n parts V_λ is also a representation of the Lie algebra $\mathfrak{gl}_{\mathbb{C}}(n)$. This representation can be in turn restricted to a representation of $\mathfrak{gl}_{\mathbb{R}}(n)$. The real Lie algebra $\mathfrak{gl}_{\mathbb{R}}(n)$ consists of all real $n \times n$ matrices and has n^2 generators ($E_{ij}, 1 \le i, j \le n$) satisfying commutation relations (2.12). These generators can be represented in terms of $n \times n$ matrices with E_{ij} being a matrix of all zeros except for the $(ij)^{\text{th}}$ entry (these are Cartan-Weyl matrices).

Now define

$$\begin{gathered} A_{ii} = iE_{ii} , \\ A_{ij} = i(E_{ij} + E_{ji}), \quad A_{ji} = E_{ji} - E_{ij} \quad \text{for } i < j . \end{gathered}$$

The operators $(A_{ij}, 1 \leq i, j \leq n)$ span $\mathfrak{u}(n)$, the Lie algebra of the Unitary group of degree n. Note that $\langle E_{k,k+1} e_m, e_{\widehat{m}} \rangle = \langle e_m, E_{k+1,k} e_{\widehat{m}} \rangle$ and so $\langle E_{ij} e_m, e_{\widehat{m}} \rangle = \langle e_m, E_{ji} e_{\widehat{m}} \rangle$ by (2.13). It follows for all $e_m, e_{\widehat{m}} \in V_\lambda$ that

$$\langle A_{ij} e_m, e_{\widehat{m}} \rangle = -\langle e_m, A_{ij} e_{\widehat{m}} \rangle \ ,$$

i.e. operators A_{ij} are anti-Hermitian in the space V_λ. By imposing the condition $tr(A) = 0$ for any $X \in \mathfrak{u}(n)$ we identify a subalgebra $\mathfrak{su}(n)$, the Lie algebra of the special Unitary group of degree n. Recall that the exponential map $\exp : \mathfrak{su}(n) \to SU(n)$ is onto, and so every $g \in SU(n)$ can be expressed as $\exp(tX)$ for some $X \in \mathfrak{su}(n)$ and $t \in \mathbb{R}$. But since each X is anti-Hermitian, i.e. $X = -X^*$, we have for any $e_m, e_{\widehat{m}} \in V_\lambda$

$$\begin{aligned}\langle g e_m, g e_{\widehat{m}} \rangle = \langle \exp(tX) e_m, \exp(tX) e_{\widehat{m}} \rangle &= \langle e_m, (\exp(tX))^* \exp(tX) e_{\widehat{m}} \rangle \\ &= \langle e_m, \exp(t(X^* + X)) e_{\widehat{m}} \rangle = \langle e_m, e_{\widehat{m}} \rangle \ ,\end{aligned}$$

where we can go from line one to line two because X and X^* commute ([22, Thm. II (a), p. 378]). In other words the group $SU(n)$ acts on V_λ unitarily. □

Finally we describe the quantisation of $U(\mathfrak{gl}(n))$. The quantum algebra $U_q(\mathfrak{gl}(n))$, $q \in (0,1)$, is a unital associative complex algebra with generators $\{E_i^\pm; 1 \leq i \leq n\}$, $\{q^{\pm \epsilon_i/2}, 1 \leq i \leq n\}$ satisfying the following relations

$$\begin{aligned}
q^{\epsilon_i/2} q^{-\epsilon_i/2} &= q^{-\epsilon_i/2} q^{\epsilon_i/2} = 1 \ , \\
q^{\epsilon_i/2} q^{\epsilon_j/2} &= q^{\epsilon_j/2} q^{\epsilon_i/2} \ , \\
q^{\epsilon_i/2} E_j^\pm q^{-\epsilon_i/2} &= q^{\pm 1/2} E_j^\pm \qquad \text{for } i = j \ , \\
&= q^{\mp 1/2} E_j^\pm \qquad \text{for } i = j+1 \ , \\
&= E_j^\pm \qquad\qquad \text{otherwise} \ , \\
[E_i^+, E_j^-] &= \delta_{ij} \frac{q^{H_i} - q^{-H_i}}{q - q^{-1}} \ , \\
(E_i^\pm)^2 E^\pm - (q + q^{-1}) E_i^\pm E_j^\pm E_i^\pm + E_j^\pm (E_i^\pm)^2 &= 0 \quad \text{for } |i - j| = 1 \ , \\
E_i^\pm E_j^\pm &= E_j^\pm E_i^\pm \qquad \text{for } |i - j| \geq 2 \ ,
\end{aligned}$$

where $H_i = \epsilon_i - \epsilon_{i+1}$.

The coproduct $\triangle : U_q \to U_q$ is given by

$$\triangle(q^{\epsilon_i/2}) = q^{\epsilon_i/2} \otimes q^{\epsilon_i/2} \ ,$$
$$\triangle(E_i^{\pm}) = E_i^{\pm} \otimes q^{-H_i/2} + q^{H_i/2} \otimes E_i^{\pm} \ .$$

Vector spaces V_λ, where λ is a partition of at most n parts, constituting all finite-dimensional representations of $U(\mathfrak{gl}(n))$ are also representations of $U_q(\mathfrak{gl}(n))$ with the same corresponding highest weight vectors. Moreover, branching rules (2.15) and (2.16) also hold. The actions of the generators on any basis vector $e_m \in V_\lambda$ are defined by

$$q^{\epsilon_i/2} e_m = q^{\left(\sum_{k=1}^{i} m_k^i - \sum_{k=1}^{i-1} m_k^{i-1}\right)/2} e_m \ ,$$
$$E_i^{+} e_m = \sum_{\widehat{m}}^{(k)} c_k(\widehat{m}, m) e_{\widehat{m}}, \quad E_i^{-} e_m = \sum_{\widehat{m}}^{(k)} c_k(m, \widehat{m}) e_{\widehat{m}} \ ,$$

where $\sum_{\widehat{m}}^{(k)}$ is defined as in (2.14b) and $c_k(m, \widehat{m})$ is given by

$$c_k(m, \widehat{m}) = \left(- \frac{\prod_{i=1}^{k-1} [l_i^{k-1} - l_j^k]_q \prod_{i=1}^{k+1} [l_i^{k+1} - l_j^k + 1]_q}{\prod_{i \neq j} [l_i^k - l_j^k]_q [l_i^k - l_j^k + 1]_q} \right)^{\frac{1}{2}}$$

if $\widehat{m}_j^k = m_j^k - 1$ for some j and $\widehat{m}_p^i = m_p^i$ for all $(i,p) \neq (k,j)$, i.e. the pattern $\widehat{m}$ is obtained from the pattern m by subtracting 1 from a single element m_j^k, and is zero otherwise.

Finally the Clebsch-Gordan formula for the embedding of any $V_{\lambda'}$ in $V_\lambda \otimes V$ in (2.16) is again given by (2.17). The corresponding Wigner coefficients $w_q((m; 1_n, ..., i_j))$ can be written as a product of the reduced Wigner coefficients $w_q^{(1)}$ and $w_q^{(2)}$ similarly to (2.18), however the latter are now given by

$$w_q^{(1)}\left(\begin{array}{ccc|c} x_1 & \cdots & x_n & i \\ y_1 & \cdots & y_{n-1} & \end{array}\right) = \left(\frac{\prod_{k\le n-1}[\widetilde{y}_k-\widetilde{x}_i]_q}{\prod_{k\le n, k\ne i}[\widetilde{x}_k-\widetilde{x}_i+1]_q}\right)^{\frac{1}{2}} , \tag{2.20}$$

$$w_q^{(2)}\left(\begin{array}{ccc|c} x_1 & \cdots & x_n & i \\ y_1 & \cdots & y_{n-1} & j \end{array}\right) = \left(\prod_{k\le n, k\ne i}\frac{[\widetilde{x}_k-\widetilde{y}_j+1]_q}{[\widetilde{x}_k-\widetilde{x}_i+1]_q}\prod_{k\le n-1, k\ne j}\frac{[\widetilde{y}_k-\widetilde{x}_i]_q}{[\widetilde{y}_k-\widetilde{y}_j]_q}\right)^{\frac{1}{2}} , \tag{2.21}$$

where $\widetilde{x}_k = x_k - k$ and $\widetilde{y}_k = y_k - k$. The Wigner coefficients of $U(\mathfrak{gl}(n))$ can be obtained from the ones above by letting q tend to 1.

Note the multiplicative structure of the Wigner coefficients which are expressed in terms of a product of reduced Wigner coefficients, both in the classical and quantum cases. This property of w will prove invaluable in Chapter 5 when we analyse a certain Markov chain evolving in the Gelfand-Cetlin cone.

Representations of $U_q(\mathfrak{gl}(n))$ at $q = 0$ and the Robinson-Schensted algorithm. Finally we explain how representations of $U_q(\mathfrak{gl}(n))$ are connected with the Robinson-Schensted correspondence. From our discussion above it follows that the algebra $U(\mathfrak{gl}(n))$ can in some respect be regarded as the limit of $U_q(\mathfrak{gl}(n))$ as $q \to 1$. At the other hand, for $q = 0$, $U_q(\mathfrak{gl}(n))$ no longer makes sense. However, we can still take the limit $q \to 0$ in the definition of the wigner coefficients w_q. For any partition λ with $\lambda_1 \ge \lambda_2 ... \ge \lambda_n$ consider the branching rule (2.16) again. Recall that for any λ', such that $\lambda \nearrow \lambda'$, the embedding of the module $V_{\lambda'}$ in $V_\lambda \otimes V$ is described by the Clebsch-Gordan formula (2.17) which gives the basis vectors of $V_{\lambda'}$ as linear combinations of the basis vectors of $V_\lambda \otimes V$. For any GC-pattern m denote by $\tau(m)$ the corresponding semi-standard Young tableau. In [24] Date, Jimbo and Miwa show ([24, Prop. 2.4]) that for any $e'_m \in V_{\lambda'}$ and $e_{\widehat{m}} \in V_\lambda$

$$\lim_{q\to 0} e'_m = \alpha e_{\widehat{m}} \otimes e_j , \tag{2.22}$$

where $\alpha \in \mathbb{C}$ is some scalar and $\tau(m)$ is the semistandard Young tableau obtained from

$\tau(\widehat{m})$ by updating it via column-insertion with a box with number j. Authors also show that the k-fold tensor product of the natural representation $V \simeq \mathbb{C}^n$ has the following irreducible decomposition

$$V^{\otimes k} \simeq \bigotimes_T V(T)\,,$$

where $V(T)$ are irreducible $U_q(\mathfrak{gl}(n))$-modules parameterised by *standard* Young tableaux with k boxes. Moreover, using (2.22) iteratively it is shown that in the limit $q \to 0$ each vector space $V(T)$ is spanned by a vector $e_{i_1} \otimes \cdots \otimes e_{i_k}$, with $(i_1, \ldots, i_k) \in [n]^k$, such that T is the recording tableau corresponding to the Robinson-Schensted algorithm with the column insertion applied to the word $(i_1, \ldots, i_k)$.

Chapter 3

Pitman's theorem and radial part of the 3-d Brownian Motion I: some discrete Markov chains

Consider a standard Brownian motion (BM) $X := (X(t); t \geq 0)$ and a Bessel process of dimension 3 (BES^3) $R := (R(t); t \geq 0)$, both started at the origin. The present chapter and Chapter 4 are motivated by an intimate relationship enjoyed by the two processes, and in particular by a desire to link the two well known couplings of X and R given by the now famous Pitman's theorem on one hand, and the characterisation of BES^3 as the radial part of a 3-dimensional Brownian motion, on the other. Our ultimate goal is to construct a one-parameter family of bivariate processes $Z^{(\theta)} = (X, R)$, $\theta \in (0, \infty)$, taking values in the 2-dimensional wedge

$$W = \{(x, r) \in \mathbb{R} \times \mathbb{R}^+ : |x| \leq r\} , \tag{3.1}$$

with the property that X is distributed as the standard Brownian motion and R as the BES^3 process and such that by letting parameter θ tend to ∞ and 0 we can recover couplings of the two processes coming from the Pitman's theorem and by considering radial part of the 3-dimensional BM respectively.

Moreover, we also study a generalisation to the non-symmetric set-up: first we describe two couplings $(\mathrm{BM}(\mu), \mathrm{BES}^3(\mu))$, where $\mathrm{BM}(\mu)$ is a Brownian motion with drift $\mu > 0$ and $\mathrm{BES}^3(\mu)$ is a so-called 3-dimensional Bessel process of drifting Brownian motion, one coupling coming from an extension of the Pitman's theorem to drifting Brownian motion and the other from looking at a 3-dimensional drifting Brownian motion and its radial part. We then construct a bivariate family of processes depending on two parameters, θ and μ, and 'interpolating' in some way between the two constructions.

In the first several introductory sections we describe various ways in which the one-dimensional Brownian motion with drift $\mu \geq 0$ and the associated three-dimensional Bessel process, $\mathrm{BES}^3(\mu)$, are connected. In addition to the above mentioned connections we describe construction of the $\mathrm{BES}^3(\mu)$-process as an h-transform of a one-dimensional drifting Brownian motion killed when it hits the origin. We will explain how these constructions fit into a larger framework of theory of conditioned and non-intersecting processes and also describe a link to the random matrix theory.

Next we discuss the discrete versions of the $(\mathrm{BM}(\mu), \mathrm{BES}^3(\mu))$ couplings mentioned above: in particular we present the discrete Pitman's theorem and the discrete analogue of the construction of the $\mathrm{BES}^3(\mu)$-process as the modulus of $\mathrm{BM}^3(\mu)$ which comes from quantum probability considerations. We then construct a family of two-dimensional discrete Markov chains interpolating between the two constructions. Finally we identify the family of bivariate diffusion $(Z^{(\theta,\mu)}; \theta \in (0,\infty), \mu \geq 0)$ of interest as the weak limit of these Markov chains under diffusive scaling.

3.1 Two classical constructions

A *3-dimensional Bessel process* $(R_t; t \geq 0)$ *started at* $r \geq 0$ is the unique strong solution to the stochastic differential equation

$$R_t = r + \beta_t + \int_0^t \frac{1}{R_s} ds, \quad t \geq 0 . \tag{3.2}$$

One of the classical constructions of the BES^3 process is by identifying it as the radial part of a 3-dimensional standard Brownian motion. Let $B = (X(t), Y(t), Z(t); t \geq$

0) be a standard 3-dimensional BM started at the origin. A straightforward application of Itô's lemma shows that the process $|B_t| = \sqrt{X_t^2 + Y_t^2 + Z_t^2}$ satisfies SDE (3.2) with $r = 0$. Moreover the joint process of R and X is a diffusion with values in W and the generator

$$\mathscr{G}^{(0)} = \frac{1}{2}\frac{\partial^2}{\partial x^2} + \frac{1}{2}\frac{\partial^2}{\partial r^2} + \frac{x}{r}\frac{\partial^2}{\partial x dr} + \frac{1}{r}\frac{\partial}{\partial r}\ . \tag{3.3}$$

We denote this bivariate process by $Z^{(0)} = (X,R)$. The boundary of W is the set $\partial W = \{(x,r) \in \mathbb{R} \times \mathbb{R}^+ : |x| = r\}$. For any time $t > 0$, $Z_t^{(0)} \in \partial W$, if and only if $(Y(t), Z(t)) = (0,0)$. However, we know that the two-dimensional Brownian motion is transient and the probability of this event is zero. The origin is the entrance point for $Z^{(0)}$: if started at 0, the process leaves it immediately and hits the boundary again with probability zero.

This construction also allows us to see that 0 is an entrance point for the BES^3 process: if started at the origin, it leaves immediately and never returns. If started away from the origin, it never hits zero.

The second coupling is given by the celebrated Pitman's theorem.

Theorem 3.1. *(Pitman [62]) Let $X = (X(t); t \geq 0)$ be a standard Brownian motion and let $R = (R(t); t \geq 0)$ be a 3-dimensional Bessel process, both started at the origin. Then*

$$(2M_t - X_t, M_t) \overset{distr}{=} (R_t, J_t)\ , \tag{3.4}$$

where $M_t = \sup_{s \leq t} X_s$ and $J_t = \inf_{s \geq t} J_s$. In particular, $(2M_t - X_t; t \geq 0)$ is a BES^3 process started at the origin.

The Pitman's transformation $X \to 2\sup X - X$, sometimes referred to as the *Pitman transform*, which maps a path in $\mathbb{R}$ to a path in $(0,\infty)$, has been greatly generalised by O'Connell and Yor [60] and by Bougerol and Jeulin [14] (see also [12]). We discuss transformations of O'Connell and Yor in Section 2 of Chapter 5. See below for a discussion of an extension of Pitman's theorem to the drifting Brownian motion and also a paper by Rogers [67], where author describes *all* diffusions X with the property that $2\max_{s \leq t} X_s - X_t$ is again a diffusion, the so called '$2M - X$ property'.

The coupling $Z^{(\infty)} := (X,R)$, where $R = 2M - X$, takes values in W and has the associated generator

$$\mathcal{G}^{(\infty)} = \frac{1}{2}\frac{\partial^2}{\partial x^2} + \frac{1}{2}\frac{\partial^2}{\partial r^2} - \frac{\partial^2}{\partial x \partial r} \tag{3.5}$$

with the boundary condition $\frac{\partial f}{\partial r}(r,r) = 0$ for all functions f in the domain of the generator. Note that, unlike the pair $Z^{(0)} = (X,R)$ of the beginning of this section, $Z^{(\infty)}$ is a degenerate diffusion which visits the boundary of the wedge W with probability 1.

One-dimensional drifting Brownian motion and its Bessel process. The *three-dimensional Bessel process of drifting Brownian motion*, denoted $\mathrm{BES}^3(\mu)$ for $\mu > 0$, is a homogeneous diffusion with the infinitesimal generator

$$\mathcal{G}^{\mu} = \frac{1}{2}\frac{\partial^2}{\partial r^2} + \mu\coth(\mu r)\frac{\partial}{\partial r} \ . \tag{3.6}$$

In many ways $\mathrm{BES}^3(\mu)$ process is related to drifting Brownian motion, $\mathrm{BM}(\mu)$, analogously to the way BES^3 is related to the standard Brownian motion. In particular

Theorem 3.2. *(Pitman and Rogers [68, Thm. 3]) Suppose $(X(t); t \geq 0)$ is a 3-dimensional Brownian motion with a drift of magnitude $\mu > 0$, started at 0. Then the radial part of X, $|X_t| := R_t$, is an $\mathbb{R}^+$-valued diffusion with generator* (3.6).

The bivariate process $(X(t), R(t); t \geq 0) := Z^{(0,\mu)}$ with values in W arising from this construction of $\mathrm{BES}^3(\mu)$ is a diffusion with the generator

$$\mathcal{G}^{(0)}_{\mu} = \frac{1}{2}\frac{\partial^2}{\partial x^2} + \frac{1}{2}\frac{\partial^2}{\partial r^2} + \frac{x}{r}\frac{\partial^2}{\partial x dr} + \frac{x\mu+1}{r}\frac{\partial}{\partial r} + \mu\frac{\partial}{\partial x} \ . \tag{3.7}$$

An analogue of Pitman's theorem reads

Theorem 3.3. *(Pitman and Rogers [68, Thm. 1]) Let $(X(t); t \geq 0)$ be a standard Brownian motion with drift $\mu \in \mathbb{R}$ started at the origin. Then*

$$R_t = 2\max_{s \leq t} X_s - X_t$$

is distributed as the radial part of a three-dimensional Brownian motion with drift of magnitude $|\mu|$.

The generator of the associated pair $(\mathrm{BM}(\mu), \mathrm{BES}^3(\mu)) := Z^{(\infty,\mu)}$, taking values in W, is given by

$$\mathscr{G}_{\mu}^{(\infty)} = \frac{1}{2}\frac{\partial^2}{\partial x^2} + \frac{1}{2}\frac{\partial^2}{\partial r^2} - \frac{\partial^2}{\partial x \partial r} + \mu\frac{\partial}{\partial x} - \mu\frac{\partial}{\partial r} \tag{3.8}$$

together with the boundary condition $\frac{\partial f}{\partial r}(r,r) = 0$ for all functions f in the domain of the generator.

h-transformed Brownian motion. Let $X = (X(t); t \geq 0)$ be a standard 1-dimensional Brownian motion started at $x > 0$ and let $\mathbb{P}_x$ be the corresponding measure. A strictly positive function $h : \mathbb{R} \to (0, \infty)$, integrable with respect to $\mathbb{P}$ is called *harmonic* with respect to $\mathscr{G}$, the generator of X, if $\mathscr{G}h = 0$. Clearly, h defined by $h(x) = x$, $x > 0$, satisfies these criteria. Let $T = \inf\{t : X_t = 0\}$ be the first time X hits the origin. Then we can define a new measure $\mathbb{Q}$ as follows

$$\mathbb{Q}_x(A) = \mathbb{E}_x\left[\frac{h(X_t)}{h(x)}\mathbf{1}_{\{A, t<T\}}\right], \quad A \in \mathscr{F}_t, x > 0\,,$$

where $\{\mathscr{F}_t\}_{t\geq 0}$ is the natural filtration of X. The process with the law $\mathbb{Q}_x$ is the *Doob-*, or *h-transform* of the Brownian motion killed when it hits the origin. Since Brownian motion is a Feller-Dynkin process, and so its transition semigroup $(P_t; t \geq 0)$ is continuous in t, we can find the generator of the new process

$$\mathscr{G}_h f(x) = \lim_{t\downarrow 0}\left(\frac{1}{t}\int \frac{h(y)}{h(x)}P_t(x,y)f(y)\,\mathbf{1}_{\{t<T\}}\,dy - f(y)\right) = \frac{1}{h(x)}\mathscr{G}(fh)(x)\,, \tag{3.9}$$

where and f is any function such that the limit above, as t tends to 0, exists. For $h(x) = x$ one easily calculates the right-hand side of the above to be

$$\mathscr{G}_r = \frac{1}{2}\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r}\,,$$

i.e. the h-transform of the Brownian motion killed when it hits the origin is distributed as the 3-dimensional Bessel process.

Moreover, we can check this directly. The density of the latter is given by

$$q_t(x,y) = \sqrt{\frac{y^3}{x}} I_{1/2}\left(\frac{xy}{t}\right) \exp\left(-\frac{x^2+y^2}{2t}\right), \quad x,y>0, t>0\,, \tag{3.10}$$

where $I_{1/2}(z) = \sqrt{\frac{2}{\pi}} \frac{\sinh(z)}{\sqrt{z}}$ is the modified Bessel function of the first kind.
Then, using the expression for the density $\widehat{p}_t$ of the transition function $\widehat{P}_t$ of the Brownian motion killed upon hitting 0 (see, for example, [65, Ch. III, Ex. 1.15]), one can verify by a direct calculation that for $x > 0$ and any continuous integrable f, $\int \frac{y}{x} \widehat{p}_t(x,y) f(y) dy = \int q_t(x,y) f(y) dy$.

The h-transform of the killed Brownian motion is often called *Brownian motion conditioned not to hit the origin*.

Now let X be a BM(μ) and $\mathbb{P}^\mu$ its associated law. Let $h(x) = 1 - e^{-2\mu x}$ for $x > 0$. The function $h(x)$ vanishes at $x = 0$ and is strictly positive for $x > 0$. Moreover, one easily checks that it is harmonic with respect to the generator of the Brownian motion with drift μ, $\frac{1}{2}\frac{\partial^2}{\partial x^2} + \mu \frac{\partial}{\partial x}$. Thus, an h-transform of the Brownian motion BM(μ) killed when it hits the origin is well-defined. Using (3.9), we verify that the generator of the new process is (3.6), i.e. the transformed process is a diffusion distributed as the three-dimensional Bessel process of drifting Brownian motion.

For more on h-transforms and Brownian motion see, for example, [65, Ch. III.3] or [69, Ch. IV.39].

3.1.1 Random matrix theory

Non-intersecting particle processes and conditioned processes have strong connections to the random matrix theory. This exciting area has seen much development in the recent years; links with non-colliding processes, random growth models and representation theory have been discovered.

Let $\mathscr{H}_n$ be the space of all $n \times n$ Hermitian matrices, i.e. if $H \in \mathscr{H}_n$, then $H = H^*$, where $*$ denotes the conjugate transpose. Let dH be the Lebesgue measure on $\mathscr{H}_n$, that is, for $H = (H_{ij})_{1 \le i,j \le n} \in \mathscr{H}_n$, $dH = \prod_i dH_{ii} \prod_{i<j} d\mathrm{Re}(H_{ij}) d\mathrm{Im}(H_{ij})$, where $\mathrm{Re}(H_{ij})$

and $\mathrm{Im}(H_{ij})$ denote the real and imaginary parts of H_{ij} respectively. The *Gaussian Unitary ensemble* (GUE) is the measure

$$d\mu_n(H) = \frac{1}{C_n} e^{\mathrm{tr} H^2} dH$$

on $\mathscr{H}_n$. Here C_n is the normalisation constant. The GUE is a *unitary invariant* measure, in that for any $n \times n$ unitary matrix U, $d\mu_n(U^* H U) = d\mu_n(H)$. Matrices drawn according to this measure are called *GUE matrices* and are one of the main objects of study in random matrix theory.

A GUE matrix is a Hermitian matrix with standard complex Gaussian entries, i.e. an $n \times n$ matrix H such that $H_{ij} = (h_{ij} + i h_{ji})/\sqrt{2} = H^*_{ji}$, for $1 \leq i < j \leq n$, and $H_{ii} = h_{ii}$, for $1 \leq i \leq n$, where h_{ij} are standard Gaussian random variables. The normalisation by $\sqrt{2}$ ensures that $\mathbb{E}[|H_{ij}|^2] = 1$ for all $i \neq j$. Since H is a Hermitian matrix, it has n real eigenvalues. Let $\lambda = (\lambda_1, ..., \lambda_n)$ be a vector of ordered eigenvalues of H with $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n$. The probability density of eigenvalues of a GUE matrix is well known and is given by

$$\mathbb{P}(d\lambda) = \frac{1}{Z_n} \triangle(\lambda)^2 \prod_{i=1}^{n} e^{-d\lambda_i^2} \ , \tag{3.11}$$

where $\triangle(\lambda)$ is the Vandermonde determinant (see (2.10)) and Z_n is the normalising constant. Thus, the distribution of each eigenvalue is governed by a Gaussian measure with the Vandermonde determinant representing the repelling force between any two eigenvalues. The measure on the GUE eigenvalues is called the *Hermite ensemble*. An important property of a measure of type (3.11) is that one can rewrite it as a determinant

$$\mathbb{P}(d\lambda) = \det[(K_n(d\lambda_i, d\lambda_j))_{1 \leq i,j \leq n}] \ ,$$

where the function $K : \mathbb{R}^2 \to \mathbb{R}$ is called the *correlation kernel* and can be written in terms of polynomials orthogonal with respect to the Gaussian measure. Point processes whose measure can be written in a form of a determinant, like above, are called *determinantal*. In particular the process of the GUE eigenvalues is determinantal with the associated *Hermite kernel*. It is precisely this determinantal structure of the eigenvalue measure which lies at the heart of the links between random matrix theory and various growth models. Of a particular interest is the limiting distribution of the largest eigenvalue, appropriately scaled, as the size of the matrix tends to infinity, the *Tracy-Widom distribution* [77]. This distribution arises as the limiting distribution of largest particles of determinantal point processes describing the tiling of a so-called Aztec diamond [46], the corner growth model [43] (also interpreted as a totally asymmetric exclusion process and a last passage percolation model), the poly-nuclear growth model [63]. The Tracy-Widom distribution of the largest eigenvalue also arises in the context of random Young tableau. In particular Baik, Deift and Johansson proved that the appropriately scaled longest increasing subsequence of a random permutation, which is equivalent to the length of the top row of a random Young tableau constructed from this permutation via the Robinson-Schensted algorithm, converges to the Tracy-Widom distribution as the length of the permutation tends to infinity [3]; see also [2] and [45]. For more details on the Gaussian Unitary ensembles see Mehta [56, Ch. 6] and for a survey of related models and results in the area – a nice paper by König [52] and references therein.

Hermitian Brownian motion. Our interest lies in the process version of a GUE matrix. Let $H(t)$ be an $n \times n$ Hermitian matrix with entries given by standard Brownian motions, i.e. $H_{ij}(t) = (B_{ij}(t) + iB_{ji}(t))/\sqrt{2} = H^*_{ji}(t)$, for $i < j$, and $H_{ii}(t) = B_{ii}(t)$, where B_{ij}'s are standard independent Brownian motions started at the origin. We call the matrix-valued process $(H(t); t \geq 0)$ *Hermitian Brownian motion*. One might now wonder what the distribution of the eigenvalues of H is? In this section we answer this question for the case $n = 2$, deferring the general n-dimensional case till Chapter 5.

Example 3.4. *Consider a* 2×2 *Hermitian Brownian motion.*

$$H(t) = \begin{pmatrix} B_{11}(t) & \frac{B_{12}(t)+iB_{21}(t)}{\sqrt{2}} \\ \frac{B_{12}(t)-iB_{21}(t)}{\sqrt{2}} & B_{22}(t) \end{pmatrix}.$$

Solving the characteristic equation det$(H(t) - \mathbb{I}\lambda) = 0$*, we find an expression for* *the matrix's two eigenvalues*

$$\lambda_{1,2}(t) = \frac{\widetilde{B}_1(t)}{\sqrt{2}} \pm \frac{1}{2}\sqrt{2\widetilde{B}_2^2(t) + 2B_{12}^2(t) + 2B_{21}^2(t)}\,,$$

where we define $\widetilde{B}_1 = (B_{11} + B_{22})/\sqrt{2}$ *and* $\widetilde{B}_2 = (B_{11} - B_{22})/\sqrt{2}$*. An application of Itô's lemma yields stochastic differential equations satisfied by the two eigenvalues*

$$\lambda_1(t) = \beta_1(t) + \int_0^t \frac{1}{\lambda_1(s) - \lambda_2(s)} ds \quad \text{and} \quad \lambda_2(t) = \beta_2(t) + \int_0^t \frac{1}{\lambda_2(s) - \lambda_1(s)} ds\,,$$

where β_1 *and* β_2 *are two independent standard Brownian motions. Moreover, the normalised difference of the eigenvalues*

$$\frac{\lambda_1(t) - \lambda_2(t)}{\sqrt{2}} = \sqrt{\widetilde{B}_2^2(t) + B_{12}^2(t) + B_{21}^2(t)}$$

is distributed as the 3-dimensional Bessel process. Thus, coupled with $\widetilde{B}_2$*, the normalised difference of the diagonal entries, this difference forms a bivariate diffusion with the generator* (3.3)*. Also, interestingly, the triplet* $(\lambda_1(t), \lambda_2(t), \mu(t); t \geq 0)$*, where* μ *is the eigenvalue of the first principal minor of* H*, i.e. the first diagonal entry of* H*, is also a diffusion. Again, by applying Itô's formula, one finds its generator to be*

$$\begin{aligned}\mathscr{G} = \frac{1}{2}\left[\frac{\partial^2}{\partial \lambda_1^2} + \frac{\partial^2}{\partial \lambda_2^2} + \frac{\partial^2}{\partial \mu^2}\right] + \frac{\partial}{\partial \lambda_1}\frac{1}{\lambda_1 - \lambda_2} + \frac{\partial}{\partial \lambda_2}\frac{1}{\lambda_2 - \lambda_1} +& \\ + \frac{\mu - \lambda_2}{\lambda_1 - \lambda_2}\frac{\partial^2}{\partial \lambda_1 \partial \mu} + \frac{\lambda_1 - \mu}{\lambda_1 - \lambda_2}\frac{\partial^2}{\partial \lambda_2 \partial \mu}&\,.\end{aligned}$$

We will look at higher-dimensional examples of joint processes of eigenvalues of GUE minors in Chapter 5.

◇

3.2 Markov functions of Markov processes

Finding bivariate Markov processes with specified Markov marginals is related to a more general question of when a measurable function of a Markov process is again Markov. Of course in general - it isn't. There are, however, various results detailing sufficient and necessary conditions ensuring that it is. The most important sufficient criteria are the so-called *Dynkin criterion* and the *intertwining condition*. In particular, intertwining plays an important role in the Pitman's theorem and its extensions.

3.2.1 Dynkin criterion and the Chapman-Kolmogorov equation

Let $X = (X(t); t \geq 0)$ be a (discrete or continuous) Markov process, defined on some probability space $(\Omega, \mathscr{F}, \mathbb{P})$, with a measurable state space $(S, \mathscr{B})$ and an initial distribution μ, i.e. $\mu(A) = \mathbb{P}(X(0) \in A)$ for any $A \in \mathscr{B}$. We define the transition function of X as usual

$$P_t(x, A) = \mathbb{P}(X(s+t) \in A | X(s) = x), \qquad \text{for all } A \in \mathscr{B}, t > 0 .$$

Let $(S', \mathscr{B}')$ be another measurable statespace and $\phi : S \to S'$ a measurable map taking S to S'. Under what conditions is $Y := \phi(X)$, a measurable function of X, again Markov? The question has been widely discussed in the literature and the simplest sufficient condition dates back to Dynkin [27].

Theorem 3.5. *(Dynkin criterion [27, Thm. 10.13]) Let the setup be as above. Suppose that for any $t > 0$ and $A \in \mathscr{B}'$, and for any $x, x' \in S$ such that $\phi(x) = \phi(x') \in S'$*

$$P_t(x, \phi^{-1}(A)) = P_t(x', \phi^{-1}(A)) . \tag{3.12}$$

Then the process $Y = (Y(t); t \geq 0)$ is Markov with state space $(S', \mathscr{B}')$ and transition probabilities $\widetilde{P}$ defined with respect to the measure

$$\mathbb{P}(Y(t) \in A) = \mathbb{P}(X(t) \in \phi^{-1}(A)), \qquad A \in \mathscr{B}', t \geq 0 .$$

We call condition (3.12) *Dynkin criterion.*

In other words, if the conditional distribution of Y_{s+t} under $\mathbb{P}$ depends on X_s only through $\phi(X_s)$, then Y is Markov for any initial distribution μ of X. In particular, if X is a discrete-time Markov chain, then (3.12) translates to

$$\mathbb{P}(Y_n = y_n | X_{n-1} = x_{n-1}) = \mathbb{P}(Y_n = y_n | \phi(X_{n-1}) = y_{n-1})$$

for all $y_n \in S'$, $x_{n-1} \in S$ and $n \geq 1$.

Note that under Dynkin condition the new process Y is not only Markov with respect to its own filtration, but also with respect to a larger filtration generated by the original process X. Three-dimensional Bessel process, viewed as the radial part of the three-dimensional Brownian motion is an example of Dynkin's criteria. One can see directly from the form of the generator (3.3) that the marginal process R, which is a function of the bi-variate process $Z^{(0)} = (X, R)$ with $\phi(x, r) = r$, is Markov. Moreover the diffusion and the drift coefficients of R do not depend on X and one concludes that in fact R is distributed as the BES^3-process.

Various papers discussed and extended Dynkin's result. In [18] Burke and Rosenblatt proved that condition (3.12) is necessary and sufficient in the cases when X is a stationary reversible chain with finite state space and invariant initial distribution or when X is a continuous-time Markov chain with finite state space and stationary transition probability functions continuous in time. Authors also give a necessary condition on the transition probability matrix P of a discrete-time X in order for $\phi(X)$ to be Markov for any measurable transformation ϕ. A paper [40] by Hachigian and Rosenblatt considers a stationary and reversible Markov chain living on a more general state space. See also [39], which extends results of [18] to a chain with a denumerable state space.

In [70] Rosenblatt relates the Markov property of the new process $\phi(X)$ to whether its first-order transition probability functions satisfy the Chapman-Kolmogorov equation. Recall that a transition probability function P of a discrete parameter chain satisfies the Chapman-Kolmogorv equation if

$$P_{n+m}(x,A) = \sum_{x'} P_n(x,x')P_m(x',A), \quad n,m \in \mathbb{N}\,, \tag{3.13}$$

or, in case X is a continuous parameter process,

$$P_{t+s}(x,A) = \int_{x'} P_t(x,dx')P_s(x',A), \quad s,t > 0 \tag{3.14}$$

for all $x \in S$, $A \in \mathscr{B}$. Transition probability functions of a Markov process necessarily satisfy the Chapman-Kolmolgorov equations. But reverse does not have to hold - one can find a non-Markovian chain whose transition probabilities satisfy (3.13) (or (3.14)). See, for example, a note by Feller [30].

Rosenblatt derives conditions on P, the transition probability functions of X, necessary for the first-order transition probability functions $\widetilde{P}$ of Y to satisfy the Chapman-Kolmogorov equation. He then proceeds to prove that these conditions are also sufficient for Y to be Markov and in fact imply the Dynkin criteria, making them a special case of the Dynkin's result. [40] and [39] also discuss the problem from the point of view of the Chapman-Kolmogorov equation.

See also a paper by Kelly [50], where author derives some Dynkin-like sufficient and necessary conditions for a function of a Markov chain to be Markov in the context of discrete-time Markov processes with countable state space.

3.2.2 Intertwining

Suppose now $X = (X_t; t \geq 0)$ is a continuous-parameter Markov process with transition functions $(P_t; t \geq 0)$. Define

$$P_t f(x) = \int_S P_t(x,dx')f(x')$$

for any measurable, bounded function f. Again let ϕ be a measurable function from S to S' and $Y = \phi(X)$. Let Λ be a Markov kernel from S' to S, that is for any $y \in S'$, $\Lambda(y,\cdot)$ is a probability measure on S and for any $A \in \mathscr{B}$, $\Lambda(\cdot,A)$ is a bounded measurable function on S'. For any measurable bounded S-valued function f we will write $\Lambda f(y)$ for the integral $\int_S \Lambda(y, dx) f(x)$. Suppose Q_t is a probability semigroup on $(S', \mathscr{B}')$. We say that P_t *and* Q_t *are intertwined with respect to the kernel* Λ if for all $y \in S'$ and $A \in \mathscr{B}$, $t \geq 0$

$$\int_S \Lambda(y, dx) P_t(x, A) = \int_{S'} Q_t(y, dy') \Lambda(y', A), \quad t \geq 0 ,$$

or more concisely

$$\Lambda P_t = Q_t \Lambda, \quad t \geq 0 . \tag{3.15}$$

We call Λ an *intertwining kernel*.

If X and Y are discrete time Markov chains with n-step transition functions denoted by P_n and Q_n respectively, then we say P and Q are intertwined with respect to some discrete Markov kernel Λ if for all $y \in S'$ and $x \in S$

$$\sum_{x'} \Lambda(y, x') P_n(x', x) = \sum_{y'} Q_n(y, y') \Lambda(y', x), \quad n \geq 1 .$$

Define a kernel Φ on bounded measurable S'-valued functions, denoted $b\mathscr{B}'$, by

$$\Phi f = f \circ \phi, \quad f \in b\mathscr{B}' .$$

The following result is due to Pitman and Rogers

Theorem 3.6. *([68, Thm. 2]) Suppose there exists a Markov kernel Λ from S' to S such that*

$$\Lambda \Phi f = f, \qquad \forall f \in b\mathscr{B}' ,$$

and the semigroup defined by $Q_t = \Lambda P_t \Phi$ satisfies the intertwining relationship (3.15). *Let X be a Markov process with initial law $\Lambda(y, \cdot)$ for some $y \in S'$ and a transition semigroup $(P_t, t \geq 0)$. Then $Y := \phi(X)$ is a Markov process with $Y_0 = y$ and a transition semigroup $(Q_t, t \geq 0)$. Moreover, the following filtering relationship holds*

$$\mathbb{P}(X_t \in A|\phi(X_s), 0 \le s \le t) = \Lambda(\phi(X_t), A) \quad a.s \tag{3.16}$$

for all $t \ge 0$ *and* $A \in \mathscr{B}$.

The result naturally extends to the set-up of discrete-time Markov chains.

The n-dimensional Bessel process, for $n \in \mathbb{N}$, viewed as the radial part, and so a function, of the n-dimensional Brownian motion, with or without drift, provides examples of intertwinings. We are in particular interested in the 3-dimensional case. Let P_t^μ and Q_t^μ be the semigroups of the three-dimensional Brownian motion with drift of magnitude $|\mu| \ge 0$ and its radial part respectively. Then, for all $t \ge 0$ [68]

$$\Lambda^\mu P_t^\mu = Q_t^\mu \Lambda^\mu \ ,$$

where, for $\mu > 0$, $\Lambda^\mu(r, \cdot)$ is the von Mises distribution on the sphere of radius r in $\mathbb{R}^3$. When $\mu = 0$, $\Lambda^0 := \Lambda$ is the uniform distribution on the (surface of the) sphere of radius r in $\mathbb{R}^3$. As a consequence it is possible to deduce Theorem 3.2.

Intertwining also plays an important role in the context of Pitman's theorem. Let $M_t := \sup_{s \le t} X_s$, where X is the drifting Brownian motion. By showing that the semigroups of the bivariate process $(U, M) := (M - X, M)$ and the process $2M - X = U + M$ are intertwined, Rogers and Pitman prove Theorem 3.3. One can find more examples of intertwinings involving Brownian motions in [10], [20], [26], and [79].

In conclusion we mention the connection between the Dynkin and the intertwining conditions. Pitman and Rogers [68] observed that if the intertwining condition holds for a pair of processes $(X, Y = \phi(X))$, then the Dynkin condition must apply to their reverses. Kelly [50], on the other hand, pointed out that, in the context of discrete-time Markov chains, if Dynkin holds for $(X, Y = \phi(X))$, then the intertwining condition holds for the reversed chains.

In the original paper [62] Pitman proved his famous result by first considering its discrete version. He constructed discrete chains approximating the standard Brownian motion and the BES^3-process and satisfying the same relationship as he hoped to prove their continuous counterparts to satisfy; taking the diffusion limit then yielded the desired result. We follow the same strategy and first identify a one-parameter family of discrete-time Markov chains with the marginals given by the simple symmetric random walk (SSRW) and the so called discrete Bessel process, dBES^3 (see below for definition). We also construct a family of discrete bivariate chains with the marginals given by a drifting random walk and the corresponding discrete Bessel process, $\mathrm{dBES}^3(\mu)$, also defined below. In section 3.7, by applying the appropriate scaling and taking the weak limit, we arrive at the family of diffusions $(Z^{(\theta,\mu)}, \theta \in [0,\infty), \mu \geq 0)$ of interest.

As a motivation we start by discussing the aforementioned discrete Pitman's theorem; we identify the result and the associated pair of discrete Markov chains by looking at the dynamics of the randomly growing Young tableaux. There is no discrete equivalent of the construction of the BES^3 process as the radial part of the three-dimensional Brownian motion in classic probability. However, this analogue can be identified as a coupling of the so called spin process with a quantum random walk in quantum probability. Our objective is to construct a family of discrete bivariate Markov chains linking these two processes in an appropriate sense; this is done in section 3.5.

3.3 Discrete Pitman's theorem via randomly growing Young tableaux (q=0)

We start this section by defining two discrete-time Markov chains of interest. By a *3-dimensional discrete Bessel process*, denoted dBES^3, we mean a time-homogenous Markov chain $(R(n); n \geq 0)$ with $R(0) = 0$ taking values in $\mathbb{N} := \{0, 1, 2, ...\}$ and having the following 1-step transition probabilities

$$P(r, r+1) = \frac{r+2}{2(r+1)}, \quad P(r, r-1) = \frac{r}{2(r+1)}\ , \tag{3.17}$$

where above and in what follows for a time-homogenous Markov chain X we write $P_n(x, x')$ for $\mathbb{P}_n(X_{m+n} = x' | X_m = x)$ and $P(x, x') := P_1(x, x')$.

There are several reasons to view dBES^3 as a discrete analogue of the BES^3

process. Firstly, under the usual diffusive scaling $R_{[mt]}/\sqrt{m}$, dBES^3 converges in distribution to the three-dimensional Bessel process (see [62, Thm. 2.6]). Secondly, we can define dBES^3 as an h-transform, with $h(x) = x+1$, of a symmetric random walk killed when it becomes negative, just like continuous Bessel process of dimension 3 can be constructed as an h-transform of a Brownian motion killed when it hits 0 (see [17]).

The BES$^3(\mu)$-process, the 3-dimensional Bessel process of drifting Brownian motion, also has a discrete equivalent. The dBES$^3(\mu)$ is a homogenous discrete-time Markov chain started at 0, with values in $\mathbb{N}$ and transition probabilities given by

$$P(r,r+1) = \frac{p_1^{r+2} - p_0^{r+2}}{p_1^{r+1} - p_0^{r+1}}, \quad P(r,r-1) = p_0 p_1 \frac{p_1^r - p_0^r}{p_1^{r+1} - p_0^{r+1}}\,, \tag{3.18}$$

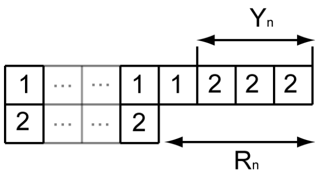
where $p_1 - p_0 = \mu$ and $p_1 + p_0 = 1$. Under the scaling $R_{[mt]}/\sqrt{m}$ with $p_1 - p_0 = \mu/\sqrt{m}$, dBES$^3(\mu)$ converges to the three-dimensional Bessel process of drifting Brownian motion BM(μ), for $\mu > 0$ (see end of section 7 of this chapter). Moreover, the Markov chain dBES$^3(\mu)$ is related to the drifting random walk in analogous ways in which the drifting Brownian motion is related to its Bessel process. Namely dBES$^3(\mu)$ is an h-transform of a drifting random walk with drift μ, killed when it becomes negative. In this case $h(r) = 1 - (p_0/p_1)^{r+1}$. Note that here the condition that $\mu > 0$, i.e. that $p_0 < p_1$, ensures that the function h is strictly positive for any $r \geq 0$. Finally, the discrete version of Theorems 3.1 and 3.3 reads

Theorem 3.7. *(Discrete Pitman's theorem) Let $X = (X_n; n \geq 0)$ be a standard random walk with $X_0 = 0$ and drift $\mu \geq 0$ and let $R = (R_n; n \geq 1)$ be a discrete BES$^3(\mu)$ process. Define $M_n := \max_{0 \leq m \leq n} X_m$ and $J_n := \inf_{m \geq n} R_m$ as the past supremum of X and the future infimum of R respectively. Then*

$$(2M - X, M) \overset{distr}{\simeq} (R, J)\,. \tag{3.19}$$

The $\mu = 0$ version of the above theorem was first discussed by Pitman [62]. Like the original theorem, discrete Pitman's construction admits extensions. In [41], for example, Hambly, Martin and O'Connell prove a version of the result for a more general class of random walks with Markovian increments, of which a simple random walk with drift is a special case. In [60] O'Connell and Yor extend the result to higher dimensional standard random walks and Poisson random walks. We discuss results from [60] in Section 2 of Chapter 5. Our treatment of the problem consists of studying the dynamics of a randomly growing Young tableau (for definition and properties see Section 2.4) and is in the spirit of [58], being effectively the $n = 2$ case of the analysis presented there.

We are going to construct a sequence of *randomly growing Young tableaux* with two rows. Let $\{\xi_n\}_{n\geq 1}$ be a sequence of independent Bernoulli random variables with mean $\mu \geq 0$ defined on some probability space $(\Omega, \mathscr{F}, \mathbb{P})$, that is $\mathbb{P}(\xi_i = 1) = p_1 = 1 - \mathbb{P}(\xi_i = -1) = 1 - p_0$, with $p_1 - p_0 = \mu$, for all $i \geq 1$, and let $X_n = \sum_{i=1}^{n} \xi_i$ with $X_0 = 0$. At time n we construct a Young tableau from a random word $(a_1, ..., a_n)$, where $a_i = 1$ if $\xi_i = 1$ and $a_i = 2$ if $\xi_i = -1$, $1 \leq i \leq n$. At time $n+1$ we update the tableau by row-inserting a_{n+1} associated to ξ_{n+1} (again see Section 2.4 for the RS algorithm with row-insertion). Thus, at each time step n we obtain a semi-standard tableau of size n with entries in $[2]^n = \{1, 2\}^n$. We are now going to identify some random variables of interest. Let R_n be the difference between the lengths of the first and the second rows. Notice that by construction R_n is always non-negative. Note that X_n, the random walk associated to the sequence of Bernoulli RV's $\{\xi_i\}_{n\geq 1}$, is just the difference between the number of 1's and 2's in the tableau at time n. We also need an auxiliary variable Y_n, which is defined to be the number of 2's in the top row, see diagram below.



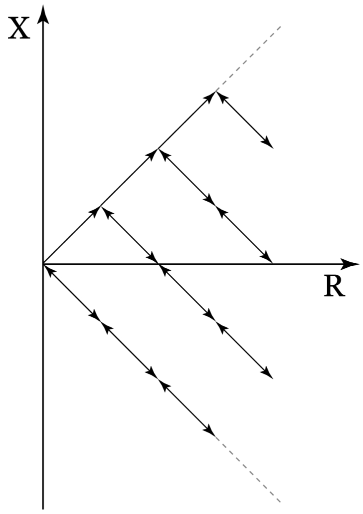
One notices that the information concerning random variables of interest can be read off from the portion of the tableau consisting of the part of the top row in excess of the bottom row. Thus, R_n is just the length of the part of the top row we are interested in, and Y_n is the number of 2's in it. Finally, for each $n \geq 1$, X_n is in fact equal to the difference between the number of 1's and 2's in the portion of the top row exceeding the bottom row; moreover, we calculate $X_n = (R_n - Y_n) - Y_n = R_n - 2Y_n$. The two tableaux below, for example, correspond to the same state of the system and we do not distinguish between the two.

1	1	1	1	1	2
2	2	2			

1	1	1	1	2
2	2			

Alternatively, we can view this as a dynamics of a *one-row* Young tableau of length R and entries in $\{1,2\}^R$.

We are interested in the pair of processes $(X_n, R_n; n \geq 0) := (Z_n^{(0,\mu)}; n \geq 0)$ whose state space is depicted below; arrows indicate possible jump directions.



One notices that by construction, given the path of the process up to time n, the value of $Z_{n+1}^{(0,\mu)}$ only depends on the current state of the system $Z_n^{(0,\mu)}$. Hence, $Z^{(0,\mu)}$ is a Markov chain. Moreover, by construction all the south-east pointing arrows in the state diagram above represent transition jumps with probability p_0 and all the north-east and north-west pointing arrows represent transition jumps with probability p_1, i.e.

$$\begin{gathered} P((x,r),(x+1,r-1)) = p_1, \quad P((x,r),(x-1,r+1)) = p_0, \quad \text{for } x < r\ , \\ P((r,r),(r+1,r+1)) = p_1, \quad \text{for } x = r\ , \end{gathered} \tag{3.20}$$

where $P((x',r'),(x,r))$ stands for $\mathbb{P}(X_{n+1} = x, R_{n+1} = r | X_n = x', R_n = r')$. In particular, if $p_1 = p_0$, then transition probability corresponding to each of the arrows in the picture above is equal to $1/2$.

Before we prove Theorem 3.7 we need to verify that R is indeed distributed as a discrete BES^3 process, which is done in the following

Lemma 3.8. *The process $(R_n; n \geq 0)$ constructed above is distributed as dBES if $\mu = 0$, and as dBES$^3(\mu)$ if $\mu \neq 0$.*

Proof. We need to show that the conditional distribution of R_{n+1} given $(R_k, 1 \leq k \leq n)$ only depends on R_n and that, specifically, the one-step transition probabilities of R are given by (3.17) if $\mu = 0$, and by (3.18) if $\mu \neq 0$. First we show that for any admissible R-path $(r_1, ..., r_n)$ (i.e. such that $|r_{i+1} - r_i| = 1$ for all $1 \leq i \leq n-1$, and $r_i \geq 0$ for all $1 \leq i \leq n$) we have

$$\mathbb{P}(R_1 = r_1, ..., R_n = r_n) = \sum_{x:(x,r_n)\in\ell_{r_n}} p_1^{\frac{n+x}{2}} p_0^{\frac{n-x}{2}}, \tag{3.21}$$

where $\ell_{r_n} = \{(x, r_n) : x \in \{-r_n, -r_n + 2, ..., r_n - 2, r_n\}\}$. Moreover, each term in the summand above is the probability of a path $((x_1, r_1), ..., (x_n, r_n))$ of the joint process, such that the marginal R-path is $(r_1, ..., r_n)$ and the endpoint $(x_n, r_n) = (x, r_n)$ belongs to ℓ_{r_n}. In particular, there are $r_n + 1$ possible Z-paths like this. We proceed by induction. Trivially $\mathbb{P}(R_1 = 1) = \mathbb{P}(X_1 = 1, R_1 = 1) + \mathbb{P}(X_1 = -1, R_1 = 1) = p_1 + p_0$. Now assume the above is true for some $n \in \mathbb{N}$ and consider an augmented R-path $(r_1, ..., r_n, r)$. Then, looking at the state diagram, we see that

$$\begin{aligned}\mathbb{P}(R_1 = r_1, ..., R_n = r_n, R_{n+1} = r_n + 1) &= \sum_{x:(x,r_n)\in\ell_{r_n}} p_1^{\frac{n+x}{2}} p_0^{\frac{n-x}{2}} p_0 + p_1^{\frac{n+r_n}{2}} p_0^{\frac{n-r_n}{2}} p_1 \\ &= \sum_{x:(x,r_n+1)\in\ell_{r_n+1}} p_1^{\frac{n+1+x}{2}} p_0^{\frac{n+1-x}{2}}\end{aligned}$$

and

$$\begin{aligned}\mathbb{P}(R_1 = r_1, ..., R_n = r_n, R_{n+1} = r_n - 1) &= \sum_{x:(x,r_n)\in\ell_{r_n}\setminus\{(r_n,r_n)\}} p_1^{\frac{n+x}{2}} p_0^{\frac{n-x}{2}} p_1 \\ &= \sum_{x:(x,r_n-1)\in\ell_{r_n-1}} p_1^{\frac{n+1+x}{2}} p_0^{\frac{n+1-x}{2}},\end{aligned}$$

which proves the induction step.

Now, for all $n \in \mathbb{N}$

$$\mathbb{P}(R_{n+1} = r | R_n = r_n, ..., R_1 = r_1) = \frac{\mathbb{P}(R_{n+1} = r, R_n = r_n, ..., R_1 = r_1)}{\mathbb{P}(R_n = r_n, ..., R_1 = r_1)} .$$

Using (3.21) the above is equal to

$$\frac{r+1}{2^{n+1}} \frac{2^n}{r_n + 1} = \frac{r+1}{2(r_n+1)}$$

if $p_1 = p_0$, and

$$\begin{aligned}
\frac{\sum_{x:(x,r_n)\in\ell_{r_n}} p_1^{\frac{n+1+x}{2}} p_0^{\frac{n+1-x}{2}}}{\sum_{x:(x,r_n)\in\ell_{r_n}} p_1^{\frac{n+x}{2}} p_0^{\frac{n-x}{2}}} &= \frac{p_1^{r_n+2} - p_0^{r_n+2}}{p_1^{r_n+1} - p_0^{r_n+1}} \quad \text{for } r = r_n + 1 \\
&= p_0 p_1 \frac{p_1^{r_n} - p_0^{r_n}}{p_1^{r_n+1} - p_0^{r_n+1}} \quad \text{for } r = r_n - 1
\end{aligned}$$

if $p_1 \neq p_0$, which concludes the proof.

□

Proof of theorem 3.7. First of all observe that $J_n := \min_{m \geq n} R_n$, the future minimum of R_n, is given by $R_n - Y_n$. To see why this is true, notice that for any $n \in \mathbb{N}$ the most that R_n can decrease by, as tableau is updated, is the number of 2's in the top row at time n which is given by Y_n (this will happen if all the 2's are row-bumped into the second row by incoming 1's). This makes $R_n - Y_n$ the smallest possible value for R_m for $m \geq n$. To see that this value is indeed attained by R, denote by $\widehat{X}$ a SSRW and note that, since X has a non-negative drift, then

$$\begin{aligned}
\mathbb{P}(R_m = R_n - Y_n = J_n,\ \text{some } m > n | J_n = j, Y_n = y) \\
= \mathbb{P}(X_m = j,\ \text{some } m \geq n | X_n = j - y) \\
\geq \mathbb{P}(\widehat{X}_m = j,\ \text{some } m > n | \widehat{X}_n = j - y) = 1 ,
\end{aligned}$$

where we have used the identity $X_n = R_n - 2Y_n$.

Next we notice that $J_n = R_n - Y_n$ is also the past maximum of X_n, denoted by M_n. Without loss of generality, suppose that $Y_n = 0$ for some $n \in \mathbb{N}$. Then $X_n = J_n := j$. If, $\xi_{n+1} = 1$, then $X_{n+1} = X_n + 1$ and $J_{n+1} = J_n + 1 = X_n + 1$. If $\xi_{n+1} = -1$, then $X_{n+1} = X_n - 1$ and $J_{n+1} = J_n = j$. Moreover, in general J can only increase if X exceeds the value j and stays constant until it happens, i.e. it is equal to the past supremum of X.

Recall that for each $n \in \mathbb{N}$ X_n is the difference between the number of 1's and 2's in the portion of the top row exceeding the bottom row. Therefore, using $X_n = R_n - 2Y_n$ again, we write

$$2M_n - X_n = 2(R_n - Y_n) - (R_n - 2Y_n) = R_n \ .$$

It is now evident that the pairs of processes $(2M - X, M)$ and (R, J) have the same one-step transition probabilities, which is enough to prove equality of laws of Markov chains. Applying Lemma 3.8 completes the proof.

□

Proof of the discrete Pitman's theorem through study of a randomly evolving Young tableau is enlightening as it hints at the connection of the problem to representation theory. As was pointed out before, at each time step $n \geq 1$ the pair (X_n, R_n) is associated to a one-row Young tableau of length R_n and a filling in $\{1,2\}^{R_n}$. Furthermore, each of the possible $R_n + 1$ fillings of such tableau is associated to a particular value of X_n. Recall that for each Young diagram with one row of length $n \geq 1$ there is an associated $(n+1)$-dimensional irreducible representation V_n of $U(\mathfrak{sl}(2))$, whose basis vectors are parameterised by the fillings of the diagram with 1's and 2's. Hence, for each $n \geq 1$ (X_n, R_n) can be associated to a basis vector of V_{R_n}. At time $n+1$ R_{n+1} jumps to either $R_n + 1$, a state associated to V_{R_n+1}, or it jumps to $R_n - 1$, a state associated to V_{R_n-1}. We thus have a Markov chain "traveling from basis vectors of one finite-dimensional space to basis vectors of another finite-dimensional space". What is more, in the symmetric case

$$P(r, r+1) = \frac{\dim(V_{r+1})}{\dim(V_r \otimes V_1)}, \quad P(r, r-1) = \frac{\dim(V_{r-1})}{\dim(V_r \otimes V_1)} \ . \tag{3.22}$$

We can develop this by employing techniques of quantum probability which will be done in the next section.

3.4 2-dimensional Markov chain from quantum probability theory ($q = 1$)

In this section we show how to construct a discrete analogue of the two-dimensional diffusion $Z^{(0,\mu)}$ with generator (3.7). We will need some basic results from the theory of quantum probability, which are presented in the following section.

3.4.1 Elements of quantum probability

In this section we shall introduce some basic principles of quantum probability and show how it compares to its classical counterpart. For a detailed account of the subject see [61, Ch. 1] or [81, Ch. 10.2] for an easy introduction. We start with a motivating example. Suppose X is a random variable on a certain probability space $(\Omega, \mathscr{F}, \mathbb{P})$ which takes n finite real values $(x_1, ..., x_n)$ with probabilities $(p_1, ..., p_n)$, $\sum_i p_i = 1$. Then the expectation of X is given by $\mathbb{E}[X] = \sum_{i=1}^n x_i p_i$, which can be rewritten as

$$\mathbb{E}_{\mathbb{P}}[X] = \mathrm{tr} \begin{pmatrix} p_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & p_n \end{pmatrix} \begin{pmatrix} x_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & x_n \end{pmatrix},$$

where tr(A) is trace of A, i.e. the sum of its diagonal entires. This gives us a motivation to think of a random variable as a diagonal $n \times n$ matrix and of its distribution as a diagonal $n \times n$ matrix with trace 1.

Let Y be another random variable with denumerable state space defined on the same probability space $(\Omega, \mathscr{F}, \mathbb{P})$. In classical probability theory XY and YX are the same random variables, that is X and Y *commute*, as do their associated matrices; in particular, they have the same distribution. Quantum probability, at the other hand, can deal with *non-commuting* random variables. In fact, we will see that classical probability is a special (commuting) case of quantum probability.

We start with some preliminary notation. Let H be a complex Hilbert space with inner product $\langle \cdot, \cdot \rangle$, s.t. for any $u, v \in H$

$$\langle u, v\rangle = u^* v = \sum_i u_i^* v_i \quad \text{and} \quad \|u\| = \langle u, u\rangle^{1/2} \ .$$

Let $\mathcal{O}(H)$ denote a collection of all self-adjoint operators from H to itself. Recall that an operator A on H is called *Hermitian* or *self-adjoint* if

$$\langle Au, v\rangle = \langle u, Av\rangle, \quad \forall u, v \in H \ .$$

In case of a finite-dimensional operator A this is equivalent to A being its own conjugate transpose: $A = A^* = \overline{A}^T$.

In quantum probability an equivalent of a random variable is a Hermitian operator, which is called an *observable*. If H has dimension $n < \infty$, then each such operator $A \in \mathcal{O}(H)$ has the following spectral decomposition

$$A = \alpha_1 P_1 + \ldots + \alpha_n P_n \ , \tag{3.23}$$

where $(\alpha_1, \ldots, \alpha_n)$ are n (not necessarily distinct) real eigenvalues of A and each P_i, for $i \in \{1, \ldots, n\}$, is an orthogonal projection on the eigenspace corresponding to the eigenvalue α_i, denoted by V_i, with $P_i P_j = 0$ for all $i \neq j$.

By an orthogonal projection on a subspace $W \subseteq H$ we mean an operator $P_W \in \mathcal{O}(H)$ with $P_W = P_W^* = P_W^2$. If W is finite-dimensional with an orthonormal basis $\{w_1, w_2, \ldots\}$, then the matrix of the operator is given by

$$P_W = \sum_i w_i w_i^* \ . \tag{3.24}$$

Thus, if the eigenspace V_i, for any $i \in \{1, \ldots, n\}$, is one-dimensional, one has $P_i = v_i v_i^*$, where v_i is the normalised (i.e. $\|v\| = 1$) eigenvector corresponding to α_i. We denote a set of all orthogonal projections on H by $\mathcal{P}(H)$.

We call the projection P_i the *event that the observable A is measured as α_i*. As pointed out by Parthasarasy in [61], one can compare (3.23) with a decomposition

$$X(\omega) = x_1 \mathbf{1}_{\{E_1\}}(\omega) + \ldots + x_n \mathbf{1}_{\{E_n\}}(\omega), \quad \omega \in \Omega ,$$

of a random variable X on probability space $(\Omega, \mathscr{F}, \mathbb{P})$ taking values $(x_1, \ldots, x_n)$, where $\mathbf{1}_{\{E_i\}}$, $i \in \{1, \ldots, n\}$, is an indicator random variable of the set $E_i = \{\omega \in \Omega : X(\omega) = x_i\}$.

A Hermitian operator ρ in $\mathscr{P}(H)$ is *positive* if $\langle \rho u, u \rangle \geq 0$, $\forall u \in H$. We call a positive operator in $\mathscr{P}(H)$ with unit trace a *density matrix*. A triplet $(H, \mathscr{P}(H), \rho)$ is called a *quantum probability space*. If H is finite-dimensional, then $(H, \mathscr{P}(H), \rho)$ is called a *finite- dimensional quantum probability space*.

Let $\rho \in \mathscr{P}(H)$ be a density matrix and let λ and u be one of its eigenvalues and the corresponding eigenvector respectively. Then $\langle \rho u, u \rangle = \lambda \|u\| > 0$, which implies that ρ has positive eigenvalues. Then we can write ρ as a sum $\sum_i p_i u_i u_i^*$, where $p_i > 0$, $i \geq 1$, with $\sum_i p_i = 1$, are eigenvalues of ρ and u_i, $i \geq 1$, are the corresponding normalised eigenvectors. If ρ is a one-dimensional projection, then we can write $\psi\psi^* = \rho$ for some $\psi \in H$ with $\|\psi\| = 1$. Vector ψ is called *a pure state*. We write ρ_ψ.

We are now ready to define probabilities on possible *measurements* of an observable $A \in \mathscr{O}(H)$. If the system is measured with respect to a density matrix $\rho = \sum_i p_i u_i u_i^*$, then

$$\mathbb{P}(A \text{ is measured as } \alpha_i \text{ with respect to } \rho) = \mathbb{P}_\rho(A \text{ is measured as } \alpha_i) := \mathrm{tr}(\rho P_i) ,$$

where P_i is the orthogonal projection on the eigenspace associated to α_i. In particular, for a pure state ψ, such that $\psi\psi^* = \rho$, we have $\mathbb{P}_\rho(A \text{ is measured as } \alpha_i) = \mathrm{tr}(\psi\psi^* P_i) = \|P_i \psi\|^2$. We can now calculate expectation of an observable A in state ρ.

$$\begin{aligned} \mathbb{E}_\rho[A] = \sum_{i=1}^n \alpha_i \mathbb{P}_\rho(A \text{ is measured as } \alpha_i) = \sum_{i=1}^n \alpha_i tr(\rho P_i) = tr(\sum_{i=1}^n \alpha_i \rho P_i) \\ = tr(\rho \sum_{i=1}^n \alpha_i P_i) = tr(\rho A) , \end{aligned}$$

where the last equality comes from the decomposition (3.23) of A.

Quantum conditioning. Just like in the classic probability theory, there is a notion of conditional probability in its quantum counterpart.

Definition 3.9. *(Quantum conditioning) Let P be an orthogonal projection, i.e. a quantum event in state ρ, and such that* $\mathrm{tr}(P\rho P) \neq 0$*. A quantum probability conditioned on P is given by the state*

$$\frac{P\rho P}{\mathrm{tr}(P\rho P)} .$$

Let A and B be two observables defined on the same quantum probability space with the state ρ, and let P_α and P_β be some events associated to A and B respectively. Then, using the above definition,

$$\mathbb{P}_\rho(B \text{ is measured as } \beta | A \text{ is measured as } \alpha) = \frac{\mathrm{tr}(P_\alpha \rho P_\alpha P_\beta)}{\mathrm{tr}(P_\alpha \rho P_\alpha)} .$$

One notices that this expression is reminiscent of the classical formula for conditional probability $\mathbb{P}(B = \beta | A = \alpha) = \mathbb{P}(A = \alpha, B = \beta)/\mathbb{P}(A = \alpha)$. Indeed, since $P_\alpha^2 = P_\alpha$, $\mathrm{tr}(P_\alpha \rho P_\alpha) = \mathrm{tr}(\rho P_\alpha^2) = \mathrm{tr}(\rho P_\alpha) = \mathbb{P}_\rho(A \text{ is measured as } \alpha)$. Thus, rearranging the above expression gives

$$\mathrm{tr}(P_\alpha \rho P_\alpha P_\beta) = \mathbb{P}_\rho(B = \beta | A = \alpha) \mathbb{P}_\rho(A = \alpha)$$

and, similarly,

$$\mathrm{tr}(P_\beta \rho P_\beta P_\alpha) = \mathbb{P}_\rho(A = \alpha | B = \beta) \mathbb{P}_\rho(B = \beta)$$

So, unlike in classical probability, the right-hand sides of the above two equations are not in general equal. However, if $AB = BA$, and so $P_\alpha P_\beta = P_\beta P_\alpha$, by using $P_\alpha^2 = P_\alpha$ and $P_\beta^2 = P_\beta$, it follows that

$$\begin{aligned} \mathrm{tr}(P_\alpha \rho P_\alpha P_\beta) &= \mathrm{tr}(\rho P_\alpha P_\beta P_\alpha) = \mathrm{tr}(\rho P_\beta P_\alpha) \\ &= \mathrm{tr}(\rho P_\alpha P_\beta) = \mathrm{tr}(\rho P_\beta P_\alpha P_\beta) = \mathrm{tr}(P_\beta \rho P_\beta P_\alpha) , \end{aligned}$$

i.e.

$$\mathbb{P}_\rho(B = \beta | A = \alpha) \mathbb{P}_\rho(A = \alpha) = \mathbb{P}_\rho(A = \alpha | B = \beta) \mathbb{P}_\rho(B = \beta) ,$$

that is

$$\mathbb{P}_\rho(A \text{ is measured as } \alpha, \textit{then } B \text{ is measured as } \beta)$$
$$= \mathbb{P}_\rho(B \text{ is measured as } \beta, \textit{then } A \text{ is measured as } \alpha)\ .$$

In other words, if observables A and B commute, then it doesn't matter in what order we measure them, something that is not in general true for non-commuting quantum random variables. It is precisely when observables commute, that they have a joint distribution in the sense of the classical probability.

3.4.2 Quantum Bernoulli random variables and quantum Bernoulli random walks

In this section we introduce quantum Bernoulli random variables and show how to construct the associated quantum Bernoulli random walks. Take H to be $\mathbb{C}^2$, then $\mathscr{O}(\mathbb{C}^2) = \mathscr{H}_2$, where $\mathscr{H}_2$ is the space of all 2×2 Hermitian matrices. We denote the space of 2×2 matrices with complex entries by $M_2(\mathbb{C})$. Consider the *Pauli matrices*

$$\sigma_x = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \tag{3.25}$$

where, as usual, $i = \sqrt{-1}$. Pauli matrices are 2×2 Hermitian matrices and so can be regarded as quantum observables. The space of all 2×2 self-adjoint matrices $\mathscr{H}_2$, which is a real sub-space of $M_2(\mathbb{C})$, is spanned by the three Pauli matrices together with the identity matrix $\mathbb{I}$. Thus, $c_1\sigma_x + c_2\sigma_y + c_3\sigma_z + c_4\mathbb{I}$, for any $c_1, c_2, c_3, c_4 \in \mathbb{R}$, is a Hermitian operator and hence a quantum observable. A sub-space of $\mathscr{H}_2$ of commuting matrices, which correspond to the 'classical' random variables, is given by all 2×2 diagonal matrices.

Let us study the Pauli matrices in more detail. The spectrum of each matrix is

$\{\pm 1\}$, and the spectral decompositions of σ_x, σ_y and σ_z are given by

$$\sigma_x = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \end{pmatrix}, \quad \sigma_y = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \begin{pmatrix} 1 & -1 \end{pmatrix}$$

and

$$\sigma_z = \frac{1}{2} \begin{pmatrix} 1 \\ i \end{pmatrix} \begin{pmatrix} 1 & -i \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ -i \end{pmatrix} \begin{pmatrix} 1 & i \end{pmatrix}$$

respectively.

Let the system be in a state corresponding to the density matrix $\rho = \begin{pmatrix} p_1 & 0 \\ 0 & p_0 \end{pmatrix}$, with $p_1 + p_0 = 1$, so that our quantum probability space is given by $(\mathbb{C}^2, \mathscr{H}_2, \rho)$. Note that conjugating Pauli matrices with Unitary matrices leaves their spectrums and commutating relationships invariant. Also, for any event P associated to σ_x, σ_y or σ_z and a unitary matrix U, we have $\mathrm{tr}(\rho U P U^*) = \mathrm{tr}(U^* \rho U P)$. Since we can always choose U which diagonalises $\rho \in \mathscr{P}(\mathbb{C}^2)$, we can assume without loss of generality that ρ is diagonal.

Now,

$$\mathbb{P}_\rho(\sigma_x \text{ is measured as } 1) = \mathrm{tr} \begin{pmatrix} p_1 & 0 \\ 0 & p_0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = p_1$$

and

$$\mathbb{P}_\rho(\sigma_x \text{ is measured as } -1) = \mathrm{tr} \begin{pmatrix} p_1 & 0 \\ 0 & p_0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = p_0 \, .$$

Similarly one calculates

$$\mathbb{P}_\rho(\sigma_y \text{ is measured as } \pm 1) = \frac{1}{2} \quad \text{and} \quad \mathbb{P}_\rho(\sigma_z \text{ is measured as } \pm 1) = \frac{1}{2}$$

In other words, σ_y and σ_z are distributed as symmetric Bernoulli random variables, while σ_x is distributed as a Bernoulli random variable with drift $p_1 - p_0$. We call observables σ_x, σ_y and σ_z *quantum Bernoulli random variables* (QBRV).

Now that we have defined a quantum Bernoulli random variable, we are going to 'sum several copies of it' to obtain a quantum Bernoulli random walk (QBRW). Before we proceed, we define what we mean by a tensor product of self-adjoint operators. Suppose H_i is a Hilbert space of dimension $d_i < \infty$, for $1 \le i \le n$ and $n \in \mathbb{N}$, and let $H = \bigotimes_{i=1}^n H_i$ (tensor product of vector spaces was defined in Section 2.1). For each $1 \le i \le n$, let T_i be a self-adjoint operator on H_i whose action is diagonalisable with eigenvalues $\{\lambda_{ij}, 1 \le j \le d_i\}$ and the corresponding eigenvectors $\{u_{ij}, 1 \le j \le d_i\}$, that is $T_i v_{ij} = \lambda_{ij} v_{ij}$, for $1 \le j \le d_i$. Then we can define a new operator $T = \otimes_{i=1}^n T_i$ on H by setting

$$Tu = T(\otimes_{i=1}^n u_i) = \otimes_{i=1}^n (T_i u_i) \tag{3.26}$$

for each vector $u = \otimes_{i=1}^n u_i \in H$, such that $u_i \in H_i$ for all $1 \le i \le n$. Then T is a self-adjoint operator on H, diagonalisable with eigenvalues $\{\prod_{i=1}^n \lambda_{ik_i}, 1 \le k_i \le d_i\}$ and the corresponding eigenvectors $\{\otimes_{i=1}^n u_{ik_i}, 1 \le k_i \le d_i\}$, i.e.

$$T(\otimes_{i=1}^n u_{ik_i}) = \otimes_{i=1}^n (T_i u_{ik_i}) = \prod_{i=1}^n \lambda_{ik_i} (\otimes_{i=1}^n u_{ik_i}) \tag{3.27}$$

for all $1 \le k_i \le d_i$ and $1 \le i \le n$.

Now, let us return to constructing a quantum Bernoulli random walk. For some $n \in \mathbb{N}$ consider a Hilbert space $\bigotimes^n \mathbb{C}^2 := (\mathbb{C}^2)^{\otimes n}$ and an algebra $\bigotimes^n M_2(\mathbb{C}) := M_2(\mathbb{C})^{\otimes n}$, which are the n-fold tensor products of $\mathbb{C}^2$ and $M_2(\mathbb{C})$ respectively. For $1 \le i \le n$, define $x_i, y_i, z_i \in M_2(\mathbb{C})^{\otimes n}$ as follows

$$\begin{aligned} x_i &= \mathbb{I} \otimes \cdots \mathbb{I} \otimes \sigma_x \otimes \mathbb{I} \cdots \otimes \mathbb{I}\,, \\ y_i &= \mathbb{I} \otimes \cdots \mathbb{I} \otimes \sigma_y \otimes \mathbb{I} \cdots \otimes \mathbb{I}\,, \\ z_i &= \mathbb{I} \otimes \cdots \mathbb{I} \otimes \sigma_z \otimes \mathbb{I} \cdots \otimes \mathbb{I}\,, \end{aligned}$$

where σ_x, σ_y and σ_z appear in the i^{th} place. These are self-adjoint operators acting on $(\mathbb{C}^2)^{\otimes n}$ as defined by (3.26). In the tensor product state $\rho^{\otimes n}$, where $\rho = \begin{pmatrix} p_1 & 0 \\ 0 & p_0 \end{pmatrix}$, y_i and z_i are distributed as symmetric Bernoulli random variables and x_i is distributed as a Bernoulli random variable with drift $p_1 - p_0$, for $1 \le i \le n$ (see proof of Theorem 3.10).

Next define for $1 \le k \le n$

$$X_k = \sum_{i=1}^{k} x_i, \quad Y_k = \sum_{i=1}^{k} y_i \quad \text{and} \quad Z_k = \sum_{i=1}^{k} z_i$$

and $X_0 = Y_0 = Z_0 = 0$. Each family $(X_k; 1 \le k \le n)$, $(Y_i; 1 \le k \le n)$ and $(Z_i; 1 \le k \le n)$ consists of commuting Hermitian operators and so for each a joint distribution exists. The operators from different families, however, do not commute.

Theorem 3.10. *In state $\rho^{\otimes n}$ $(X_k; 1 \le k \le n)$ is distributed as a simple random walk with jump probabilities $\mathbb{P}(X_{k+1} = x+1 | X_k = x) = p_1$ and $\mathbb{P}(X_{k+1} = x-1 | X_k = x) = p_0 = 1 - p_1$, and $(Y_k; 1 \le k \le n)$ and $(Z_k; 1 \le k \le n)$ are distributed as simple symmetric random walks.*

Proof. We are considering the system in the quantum probability space $((\mathbb{C}^2)^{\otimes n}, \mathscr{H}_2^{\otimes n}, \rho^{\otimes n})$. Note that for any $1 \le k \le n$ we have

$$X_k - X_{k-1} = x_k, \quad Y_k - Y_{k-1} = y_k, \quad Z_k - Z_{k-1} = z_k \ .$$

Hence, to prove the theorem, it suffices to show the claim made above, that for all $1 \le k \le n$ y_k and z_k are distributed as symmetric QBRV and x_k as a biased QBRV with mean $p_1 - p_0 = \mu$. We complete the proof by proving that x_k's (resp. y_k's, z_k's) are all mutually independent for all $1 \le k \le n$.

We start by considering x_i for some $1 \le i \le n$. Recall that the spectrum of σ_x is $\{\pm 1\}$ and that $\mathbb{I}$ (a 2×2 identity matrix) has a repeated root 1 with the corresponding eigenspace spanned by $\{e_0, e_1\}$, the canonical basis of $\mathbb{C}^2$. Using our preceding discussion and expression (3.27) in particular, we conclude that x_i has eigenvalues $\{1, -1\}$, both with multiplicity 2^{n-1} and orthonormal eigenvectors

$$e_{i_1} \otimes \ldots \otimes e_{i_{k-1}} \otimes u_1 \otimes e_{i_{k+1}} \otimes \ldots \otimes e_{i_n} \ ,$$
$$e_{i_1} \otimes \ldots \otimes e_{i_{k-1}} \otimes u_0 \otimes e_{i_{k+1}} \otimes \ldots \otimes e_{i_n} \ ,$$

respectively, where (u_1, u_0) are normalised eigenvectors or σ_x corresponding to eigenvalues $\{1, -1\}$, and indices $(i_1, ..., i_{k-1}, i_{k+1}, ..., i_n)$, with $i_k \in \{0, 1\}$, run over all possible 2^{n-1} combinations of 0's and 1's. The orthogonal projections corresponding to the events that in state $\rho^{\otimes n}$ the observable x_k is measured as 1 or -1 are given by

$$P_1 = \sum \left(e_{i_1} \otimes ... \otimes e_{i_{k-1}} \otimes u_1 \otimes e_{i_{k+1}} \otimes ... \otimes e_{i_n}\right) \left(e_{i_1} \otimes ... \otimes e_{i_{k-1}} \otimes u_1 \otimes e_{i_{k+1}} \otimes ... \otimes e_{i_n}\right)^* ,$$
$$P_0 = \sum \left(e_{i_1} \otimes ... \otimes e_{i_{k-1}} \otimes u_0 \otimes e_{i_{k+1}} \otimes ... \otimes e_{i_n}\right) \left(e_{i_1} \otimes ... \otimes e_{i_{k-1}} \otimes u_0 \otimes e_{i_{k+1}} \otimes ... \otimes e_{i_n}\right)^* ,$$

respectively, where the sum is over all possible combinations of $(e_{i_1}, ..., e_{i_{k-1}}, e_{i_{k+1}}, ..., e_{i_n})$ with $i_k \in \{0, 1\}$. Now, using conjugate transpose distributivity $(A \otimes B)^* = A^* \otimes B^*$, the above becomes

$$P_1 = \sum (e_{i_1} e_{i_1}^*) \otimes ... \otimes (e_{i_{k-1}} e_{i_{k-1}}^*) \otimes (u_1 u_1^*) \otimes (e_{i_{k+1}} e_{i_{k+1}}^*) \otimes ...(e_{i_n} e_{i_n}^*) ,$$
$$P_0 = \sum (e_{i_1} e_{i_1}^*) \otimes ... \otimes (e_{i_{k-1}} e_{i_{k-1}}^*) \otimes (u_0 u_0^*) \otimes (e_{i_{k+1}} e_{i_{k+1}}^*) \otimes ...(e_{i_n} e_{i_n}^*) ,$$

And, hence, using (3.26), we obtain

$$\mathbb{P}_{\rho^{\otimes n}}(x_k \text{ is measured as } 1) = \mathrm{tr}(\rho^{\otimes n} P_1) = \mathrm{tr}\left(\sum (\rho e_{i_1} e_{i_1}^*) \otimes ... \otimes (\rho u_1 u_1^*) \otimes ...(\rho e_{i_n} e_{i_n}^*)\right) ,$$
$$\mathbb{P}_{\rho^{\otimes n}}(x_k \text{ is measured as } 0) = \mathrm{tr}(\rho^{\otimes n} P_0) = \mathrm{tr}\left(\sum (\rho e_{i_1} e_{i_1}^*) \otimes ... \otimes (\rho u_0 u_0^*) \otimes ...(\rho e_{i_n} e_{i_n}^*)\right) .$$

Using $\mathrm{tr}(A \otimes B) = \mathrm{tr}A\mathrm{tr}B$ and also $\mathrm{tr}(\rho u_1 u_1^*) = \mathrm{tr}(\rho e_1 e_1^*) = p_1$ and $\mathrm{tr}(\rho u_0 u_0^*) = \mathrm{tr}(\rho e_0 e_0^*) = p_0$, we have

$$\begin{aligned}\mathbb{P}_{\rho^{\otimes n}}(x_k = 1) &= \mathrm{tr}(\rho u_1 u_1^*) \sum \mathrm{tr}(\rho e_{i_1} e_{i_1}^*) ... \mathrm{tr}(\rho e_{i_{k-1}} e_{i_{k-1}}^*) \mathrm{tr}(\rho e_{i_{k+1}} e_{i_{k+1}}^*) ... \mathrm{tr}(\rho e_{i_n} e_{i_n}^*) \\ &= p_1 \left(p_1^{n-1} + \mathrm{C}_1^{n-1} p_1^{n-2} p_0 + ... + \mathrm{C}_1^{n-1} p_1 p_0^{n-2} + p_0^{n-1}\right) = p_1 (p_1 + p_0)^{n-1} = p_1 ,\end{aligned}$$

where we write C_k^n for $\frac{n!}{k!(n-k)!}$. Similarly one calculates

$$\mathbb{P}_{\rho^{\otimes n}}(x_k \text{ is measured as } -1) = p_0 .$$

Thus, the increments $(X_k - X_{k-1}; 1 \le k \le n)$ are distributed as quantum Bernoulli random variables with mean $p_1 - p_0$. To prove that the increments are indeed mutually independent, we use quantum conditioning discussed before: using definition 3.9 and techniques we have used above, one calculates that $\mathbb{P}_\rho(X_m - X_{m-1} = 1 | X_k - X_{k-1} = \pm 1) = p_1$ and $\mathbb{P}_\rho(X_m - X_{m-1} = -1 | X_k - X_{k-1} = \pm 1) = p_0$, for all $1 \leq k < m \leq n$. It follows that $(X_k; 1 \leq k \leq n)$ is distributed as a Bernoulli random walk with mean $p_1 - p_0$. In a similar manner one proves analogous statements about $(Y_k; 1 \leq k \leq n)$ and $(Z_k; 1 \leq k \leq n)$.

□

3.4.3 The Spin process

In this section we introduce a process, which is constructed using QBRWs. We define the *Spin process* as the family of operators

$$S_n^2 = X_n^2 + Y_n^2 + Z_n^2, \quad n \geq 1 \tag{3.28}$$

with $S_0^2 = 0$. One checks that $(S_n^2, n \geq 1)$ forms a family of commuting Hermitian operators, and so does $(S_n, n \geq 1)$ ([9, Prop. 2]), thus defining a classical stochastic process. Moreover, ([9, Prop. 2])

$$[S_n, X_m] = [S_n, Y_m] = [S_n, Z_m] = 0; \quad n \leq m$$

Theorem 3.11. *([9, Thm. 1]) The square root of the spin process $(S_k; 1 \leq k \leq n)$ with $S_0 = 0$ is a Markov chain taking values in $\mathbb{N}$, with one-step transition probabilities given by*

$$\mathbb{P}(S_{k+1} = s+1 | S_k = s) = \frac{s+2}{2(s+1)}, \quad \mathbb{P}(S_{k+1} = s-1 | S_k = s) = \frac{s}{2(s+1)}$$

if $p_0 = p_1 = 1/2$ and

$$\mathbb{P}(S_{k+1} = s+1 | S_k = s) = \frac{p_1^{s+2} - p_0^{s+2}}{p_1^{s+1} - p_0^{s+1}}, \quad \mathbb{P}(S_{k+1} = s-1 | S_k = s) = p_1 p_0 \frac{p_1^s - p_0^s}{p_1^{s+1} - p_0^{s+1}}$$

if $p_1 \neq p_0$, $0 \leq k \leq n-1$.

So, Theorem 3.11 tells us that the square root of the spin process is distributed as dBES^3, when $p_1 = p_0 = 1/2$, and as $\mathrm{dBES}^3(\mu)$, when $p_1 \neq p_0$. In [9] Biane also states and proves several results relating the spin process S and the QBRW X. For example, conditioned on X, Y and Z are distributed as symmetric random walks. On the other hand, in the symmetric case $p_1 = p_0$, in the state $\rho^{\otimes n}(\cdot|S_n = s)$, X_n is distributed uniformly on $\{-s, -s+2, ..., s-2, s\}$ and in the state $\rho^{\otimes n}(\cdot|X_n = x)$, $(X_k; 0 \leq k \leq n)$ is distributed as a SSRW conditioned to satisfy $X_n = x$. Note also that the definition of the spin process (3.28) is reminiscent of the characterisation of BES^3 and $\mathrm{BES}^3(\mu)$ as the radial part of the 3-d Brownian motion and the 3-d Brownian motion with drift respectively. In this context, a natural question arises: what is the joint distribution of the pairing (X, S) and is it connected to the diffusion with the generator (3.7)?

3.4.4 2-dimensional Markov chain associated to QBRW

In this section we construct a two-dimensional Markov chain $Z^{(1,\mu)}$, such that the marginals are given by the QBRW, X, and the square root of the spin process. An example of another such process is the discrete Pitman's construction of section 3.3. In the present section we will construct a bivariate Markov chain with the same marginals by studying finite-dimensional representations of $U(\mathfrak{sl}(2))$. Then by considering a quantum distortion of $U_q(\mathfrak{sl}(2))$ we find a whole one-parameter family of two-dimensional chains with the $(\mathrm{RW}(\mu), \mathrm{dBES}^3(\mu))$, $\mu \geq 0$, marginals. By letting the parameter q tend to 0 we reconstruct the 2-dimensional chain of the discrete Pitman's theorem.

We have to explain the construction of the bivariate process $Z^{(1,\mu)}$ with the (X, S) marginals very carefully because, since S_n does not commute with any X_m for $m < n$, there is no well-defined classical joint distribution of $(X_1, S_1 X_2, S_2, ..., X_n, S_n)$. On the other hand, $(X_n, S_n, S_{n-1}, ..., S_1)$ is a collection of commuting self-adjoint operators which, therefore, have a well-defined joint distribution in the sense of classical probability for any $n \geq 1$. Hence, by conditioning on the event $\{X_n = x, S_n = s, S_{n-1} =$

$s_{n-1}, ..., S_1 = s_1\}$, $k \geq 0$, we can calculate probabilities of the type

$$\begin{aligned}\mathbb{P}_{\rho^{\otimes n}}(X_{m+1} = x', S_{m+1} = s' | X_m = x, S_m = s, S_{m-1} = s_{m-1}, ..., S_1 = s_1) \\ = \mathbb{P}_{\rho^{\otimes n}}(X_{m+1} = x', S_{m+1} = s' | X_m = x, S_m = s), \quad k \geq 0 ,\end{aligned}$$

which is enough to construct a Markov chain.

Proposition 3.12. *In state $\rho^{\otimes n}$ we have*

$$\mathbb{P}_{\rho^{\otimes n}}(X_{m+1} = x', S_{m+1} = s' | X_m = x, S_m = s) = p_0 |\langle e_k^s \otimes e_0, e_{k'}^{s'} \rangle|^2 + p_1 |\langle e_k^s \otimes e_1, e_{k'}^{s'} \rangle|^2 \tag{3.29}$$

for $1 \leq m \leq n-1$, $s' = s \pm 1$, $x' \in \{-s', -s'+2, ..., s'-2, s'\}$ and where $x = 2k - s$ and $x' = 2k' - s'$. Here e_k^s is a unit vector spanning a subspace of $(\mathbb{C}^2)^{\otimes n}$, a common eigenspace of X_m and $(S_1, ..., S_m)$ corresponding to the eigenvalues $2k-s$ and $(s_1, ..., s_m = s)$ respectively, where $(s_1, ..., s_m)$ is any admissible path of S up to the time m ending in s; $e_{k'}^{s'}$ is defined analogously.

Proof. First of all one easily verifies that the Pauli matrices satisfy the commutation relationships

$$[\sigma_x, \sigma_y] = 2i\sigma_z, \quad [\sigma_x, \sigma_z] = -2i\sigma_y, \quad [\sigma_y, \sigma_z] = 2i\sigma_x ,$$

from which one deduces that, for all $n, m \geq 1$,

$$[X_n, Y_m] = 2iZ_{n\wedge m}, \quad [X_n, Z_m] = -2iY_{n\wedge m}, \quad [Y_n, Z_m] = 2iX_{n\wedge m} .$$

Moreover, for all $m \geq 1$, the three operators X_m, $\frac{1}{2}(Y_m + iZ_m)$ and $\frac{1}{2}(Y_m - iZ_m)$ satisfy the same commutation relationships as the three generators $\{H, X, Y\}$ of $U(\mathfrak{sl}(2))$ (or $\mathfrak{sl}(2)$) (see section 2 of Chapter 2). Therefore, operators $\{X_m, Y_m, Z_m\}$ must span a complex lie algebra isomorphic to $U(\mathfrak{sl}(2))$, and in particular their actions on the $U(\mathfrak{sl}(2))$-modules V_k, $k \geq 0$, are well known (see eqns. (2.2)).

In his version of proof of Theorem 3.11 von Waldenfels [78] points out that for any $m \geq 1$

$$(\mathbb{C}^2)^{\otimes m} \simeq \bigoplus_{\Gamma} V_\Gamma \ ,$$

where the direct product runs over all admissible paths $\Gamma = (0 = s_0, s_1, ..., s_m)$ of S up to time m, i.e. V_Γ is a common eigenspace for the operators $S_1, ..., S_m$:

$$S_i e = s_i e, \quad 1 \leq i \leq m$$

for all $e \in V_\Gamma$. For each $\Gamma = (0 = s_0, s_1, ..., s_{m-1}, s_m := s)$ V_Γ is isomorphic to the $U(\mathfrak{sl}(2))$-module V_s [78, Prop. 1]. Thus, since X_m plays the role of operator H (see equations (2.2) again), the action of X_m on $V_\Gamma \simeq V_s$ is diagonalasible with eigenvalues $\{-s, -s+2, ..., s-2, s\}$ and the corresponding eigenvectors $\{e_k^s; 0 \leq k \leq s\}$, $x = 2k - s$. Recall that the eigenvectors $\{e_k^s; 0 \leq k \leq s\}$ form an orthonormal basis for V_s. We conclude that any basis vector $e_k^s \in V_s$ spans a common eigenspace of operators X_m and $(S_1, ..., S_m)$ with the corresponding eigenvalues $2k - s$ and $(s_1, ..., s_{n-1}, s_m = s)$. In the state $\rho^{\otimes n}$ the eigenspace corresponding to the event $\{X_m = x, S_m = s, S_{n-1} = s_{m-1}, ..., S_1 = s_1, S_0 = 0\}$ is spanned by the vectors $e_k^s \otimes e_{i_{m+1}} \cdots \otimes e_{i_n}$ with indices $(i_{m+1}, ..., i_n)$ ranging over all possible 2^{n-m} combinations of 0's and 1's. Thus, the orthogonal projection corresponding to the event we will be conditioning on is given by

$$\pi := \sum_{(i_{m+1},...,i_n)} (e_k^s \otimes e_{i_{m+1}} \cdots \otimes e_{i_n})(e_k^s \otimes e_{i_{m+1}} \cdots \otimes e_{i_n})^* \ ,$$

where, as before, the summation runs over all combinations of indices $(i_{m+1}, ..., i_n)$ with $i_k \in \{0, 1\}$, for $m + 1 \leq k \leq n$, and $\{e_0, e_1\}$ is the canonical basis for $\mathbb{C}^2$.

Now, using the definition of quantum conditioning, we can write down the state corresponding to conditioning on the event $(X_m = x, S_m = s, S_{m-1} = s_{m-1}, ..., S_1 = s_1, S_0 = 0)$

$$\rho^{\otimes n}(\cdot | X_m = x, S_m = s, S_{m-1} = s_{m-1}, ..., S_1 = s_1, S_0 = 0) = \frac{\pi \rho^{\otimes n} \pi}{tr[\pi \rho^{\otimes n} \pi]} \ .$$

Let us calculate the denominator of the above fraction. Using $(A_1 \otimes B_1)(A_2 \otimes$

$B_2) = (A_1A_2) \otimes (B_1B_2)$, $\mathrm{tr}(A\otimes B) = \mathrm{tr}(A)\mathrm{tr}(B)$ and the fact that $\mathrm{tr}(\rho e_1 e_1^*) = p_1$ and $\mathrm{tr}(\rho e_0 e_0^*) = p_0$ yields

$$\begin{aligned}
\mathrm{tr}[\pi\rho^{\otimes n}\pi] &= \mathrm{tr}[\rho^{\otimes n}\pi^2] \\
&= \mathrm{tr}[\rho^{\otimes n}\pi] = \sum_{(i_{m+1},\ldots,i_n)} \mathrm{tr}[\rho^{\otimes m} e_k^s e_k^{*s}]\mathrm{tr}[\rho e_{i_{m+1}} e_{i_{m+1}}^*]\ldots\mathrm{tr}[\rho e_{i_n} e_{i_n}^*] \\
&= \mathrm{tr}[\rho^{\otimes m} e_k^s e_k^{*s}]\left(p_1^{n-m} + \mathrm{C}_1^{n-m} p_1^{n-m-1} p_0 + \ldots + \mathrm{C}_1^{n-m} p_1 p_0^{n-m-1} + p_0^{n-m}\right) \\
&= \mathrm{tr}[\rho^{\otimes m} e_k^s e_k^{*s}](p_1+p_0)^{n-m} = \mathrm{tr}[\rho^{\otimes m} e_k^s e_k^{*s}] := \tilde{p}\ .
\end{aligned}$$

In the above we have used the fact that π, being a quantum event, is an orthonormal projection and so $\pi^2 = \pi$. Now, an operator corresponding to the event of interest $\{X_{m+1} = x', S_{m+1} = s'\}$ (which is to be conditioned on π) is given by

$$P := \sum_{(i_{m+2},\ldots,i_n)} (e_{k'}^{s'} \otimes e_{i_{m+2}} \otimes \ldots \otimes e_{i_n})(e_{k'}^{s'} \otimes e_{i_{m+2}} \otimes \ldots \otimes e_{i_n})^*\ ,$$

where the sum runs over all possible combinations of indices $(i_{m+2},\ldots,i_n)$ of 0's and 1's, and $e_{k'}^{s'} \in V_{s'}$ is an eigenvector spanning a common eigenspace of X_{m+1} and S_{m+1} corresponding to eigenvalues $x = 2k' - s'$ and s' respectively. Then, writing $e_k^s \otimes e_1 = u_1$ and $e_k^s \otimes e_0 = u_0$ and using $\mathrm{tr}[e_i e_i \rho e_j e_j] = 0$ if $i \neq j$, we get

$$\begin{aligned}
&\mathbb{P}_{\rho^{\otimes n}}[X_{m+1} = x', S_{m+1} = s' | X_m = x, S_m = s, \ldots, S_1 = s_1, S_0 = s_0] = \frac{\mathrm{tr}[\pi\rho^{\otimes n}\pi P]}{tr[\pi\rho^{\otimes n}\pi]} \\
&\quad = \frac{1}{\tilde{p}} \sum_{(i_{m+2},\ldots,i_n)} \mathrm{tr}[u_1u_1^*\rho^{\otimes(m+1)}u_1u_1^* e_{k'}^{s'} e_{k'}^{*s'} + u_0u_0^*\rho^{\otimes(m+1)}u_0u_0^* e_{k'}^{s'} e_{k'}^{*s'}]\times \\
&\qquad\qquad \mathrm{tr}[e_{i_{m+2}} e_{i_{m+2}}^* \rho (e_{i_{m+2}} e_{i_{m+2}}^*)^2]\ldots\mathrm{tr}[e_{i_n} e_{i_n}^* \rho (e_{i_n} e_{i_n}^*)^2] \\
&\quad = \frac{1}{\tilde{p}} \mathrm{tr}[u_1u_1^*\rho^{\otimes(m+1)}u_1u_1^* e_{k'}^{s'} e_{k'}^{*s'} + u_0u_0^*\rho^{\otimes(m+1)}u_0u_0^* e_{k'}^{s'} e_{k'}^{*s'}](p_1+p_0)^{n-m-1} \\
&\quad = \frac{1}{\tilde{p}} \mathrm{tr}[u_1u_1^*\rho^{\otimes(m+1)}|\langle e_k^r \otimes e_1, e_{k'}^{r'}\rangle|^2 + u_0u_0^*\rho^{\otimes(m+1)}|\langle e_k^r \otimes e_0, e_{k'}^{r'}\rangle|^2]\ . \quad (3.30)
\end{aligned}$$

But

$$\mathrm{tr}[u_1u_1^*\rho^{\otimes(m+1)}] = \mathrm{tr}[(\rho^{\otimes m} e_k^s e_k^{*s}) \otimes (\rho e_1 e_1^*)] = \mathrm{tr}[\rho^{\otimes m} e_k^s e_k^{*s}] tr[\rho e_1 e_1^*] = \tilde{p} p_1\ .$$

And, similarly, $tr[u_0 u_0^* \rho^{\otimes(m+1)}] = \tilde{p} q_0$. Substituting this into (3.30) gives

$$\mathbb{P}_{\rho^{\otimes n}}[X_{m+1} = x', S_{m+1} = s' | X_m = x, S_m = s] = p_1 |\langle e_k^s \otimes e_1, e_{k'}^{s'} \rangle|^2 + p_0 |\langle e_k^s \otimes e_0, e_{k'}^{s'} \rangle|^2 \,.$$

□

Notice that the inner products featuring in the formula of Proposition 3.12 are exactly the $U(\mathfrak{sl}(2))$-Wigner coefficients we have discussed in Chapter 2. Even though, as has been mentioned before, the families of operators $(S_n; n \geq 1)$ and $(X_n; n \geq 1)$ do not commute and so a joint process (X, S) is not well-defined in the sense of classical probabilities, the one-step transition probabilities found in Proposition 3.12 can still be used to *construct* a classical Markov chain.

Definition 3.13. *The bivariate process $Z^{(1,\mu)} = (X_n, R_n; n \geq 0)$ is a Markov chain with the statespace $\boldsymbol{W} := \{(x, r) \in \mathbb{Z} \times \mathbb{N} : x \in \{-r, -r+2, ..., r-2, r\}\}$ and the following transition probabilities*

$$\begin{aligned} P((x,r),(x+1,r+1)) &= p_1 \frac{r+x+2}{2(r+1)} \,, \\ P((x,r),(x-1,r+1)) &= p_0 \frac{r-x+2}{2(r+1)} \,, \\ P((x,r),(x+1,r-1)) &= p_1 \frac{r-x}{2(r+1)} \,, \\ P((x,r),(x-1,r-1)) &= p_0 \frac{r+x}{2(r+1)} \,. \end{aligned} \tag{3.31}$$

We have obtained the above transition probabilities by substituting values of Wigner coefficients from Proposition 2.6 of Chapter 2 to (3.29).

We point out that the chain with the above transition probabilities was first constructed by Biane in [11].

3.5 q-generalisation

Finally we can define a family of bivariate Markov chains 'interpolating' between $Z^{(0,\mu)}$ and $Z^{(1,\mu)}$. Unlike in the 'extreme' cases, there isn't a couple (X, R) of naturally occurring processes corresponding to $Z^{(q,\mu)}$, $q \in (0,1)$. Rather, we can use one-step transition probabilities (3.29) again to *define* a process $Z^{(q,\mu)} = (X,R)$ for each $q \in (0,1)$. In Chapter 2 we discussed the q-deformation $U_q(\mathfrak{sl}(2))$ of the enveloping algebra $U(\mathfrak{sl}(2))$. In particular in Proposition 2.7 we calculated the Wigner coefficients describing the branching rule of tensor products of irreducible representations of $U_q(\mathfrak{sl}(2))$. Substituting these Wigner coefficients in to (3.29) we arrive at

Definition 3.14. *The bivariate process $Z^{(q,\mu)} = (X_n, R_n; n \geq 0)$ is a Markov chain with the statespace $\boldsymbol{W} = \{(x,r) \in \mathbb{Z} \times \mathbb{N} : x \in \{-r, -r+2, ..., r-2, r\}\}$ and the following transition probabilities*

$$\begin{aligned}
P((x,r),(x+1,r+1)) &= p_1 \frac{q^{x+1} - q^{-(r+1)}}{q^{r+1} - q^{-(r+1)}} \,, \\
P((x,r),(x-1,r+1)) &= p_0 \frac{q^{r+1} - q^{x-1}}{q^{r+1} - q^{-(r+1)}} \,, \\
P((x,r),(x+1,r-1)) &= p_1 \frac{q^{r+1} - q^{x+1}}{q^{r+1} - q^{-(r+1)}} \,, \\
P((x,r),(x-1,r-1)) &= p_0 \frac{q^{x-1} - q^{-(r+1)}}{q^{r+1} - q^{-(r+1)}} \,.
\end{aligned} \tag{3.32}$$

We thus obtain a family of Markov chains with values in **W**, parameterised by the number $q \in (0,1)$. Since $U(\mathfrak{sl}(2))$-Wigner coefficients can be obtained as a limit of $U_q(\mathfrak{sl}(2))$-Wigner coefficients as q tends to 1, one recovers transition probabilities (3.31) by letting $q \to 1$ in the transition probabilities above. We say that $Z^{(q,\mu)}$ is a *q-deformation of* $Z^{(1,\mu)}$. Moreover, by letting q tend to 0, one obtains a coupling $(X,R) = (\mathrm{RW}(\mu), \mathrm{dBES}^3(\mu))$ with $R_n = X_n - 2\inf_{m \leq n} X_m$, $n \geq 1$, which is an equivalent representation of the discrete Pitman's theorem.

Again, a chain with transition probabilities (3.32) was first considered by Biane [11].

Finally, it would be interesting to see if it is possible to identify the quantum 'q-equivalents' of the QBRW X and the Spin process S^2, i.e. to see whether there in fact exists a naturally occurring two-dimensional Markov chain which transition probabilities are given by (3.32).

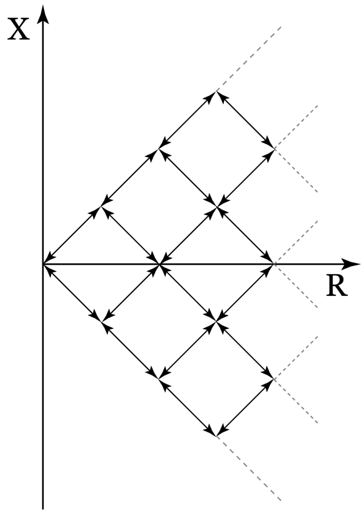
3.6 Analysing the Markov chains

We now take up the study of the one-parameter family of Markov chains found in the previous section. In particular, we calculate the distributions of the marginal processes and show that they are Markov and find some hidden intertwining relationships.

Recall that $Z^{(1,\mu)}$ is the bi-variate Markov chain with transition probabilities (3.31) given by the squares of the Wigner coefficients of $U(\mathfrak{sl}(2))$ and $Z^{(q,\mu)} := Z := (X_n, R_n; n \geq 0)$ is its q-deformed analogue (3.32); the process $Z^{(0,\mu)}$ is the chain associated to the discrete Pitman's construction (3.20). Note that for simplicity we suppress the dependence on q in the notation; a specific value of q will be explicitly indicated when it is important. Otherwise one might assume that $q \in (0,1)$. The state space for all the chains is the lattice wedge

$$\mathbf{W} = \{(x,r) \in \mathbb{Z} \times \mathbb{N} : x \in \{-r, -r+2, ..., r-2, r\}\} = \bigcup_{r \geq 0} \ell_r \,, \tag{3.33}$$

where $\ell_r = \{(x,r) : x \in \{-r, -r+2, ..., r-2, r\}\}$ and $\ell_0 = \{(0,0)\}$; see diagram below.



By $\mathscr{B}_{\mathbf{W}}$ we denote the Borel σ-algebra of subsets of $\mathbf{W}$ and $(\Omega, \mathscr{F}, \mathbb{P})$ is the underlying probability space.

3.6.1 Marginal processes

Recall that we have identified the marginal R-process of $Z^{(0,\mu)}$ as $\mathrm{dBES}^3(\mu)$, for $\mu \geq 0$, in Lemma 3.8, while the marginal processes of $Z^{(1,\mu)}$ are given by $(\mathrm{RW}(\mu), \mathrm{dBES}^3(\mu))$, for $\mu \geq 0$, by construction. In this section we prove the analogous result for $Z^{(q,\mu)}$ for a general parameter $q \in (0,1)$.

First note that for both sets of bivariate transition probabilities (3.31) and (3.32) we have

$$P((x,r),(x+1,r+1)) + P((x,r),(x+1,r-1)) = p_1 = P((r,r),(r+1,r+1))$$

and

$$P((x,r),(x-1,r+1)) + P((x,r),(x-1,r-1)) = p_0 = P((-r,r),(-r-1,r+1))$$

for all $(x,r) \in \mathbf{W}$. Hence, for $p_1 \neq p_0$ $X = (X_n; n \geq 0)$ is a simple random walk with drift $\mu := p_1 - p_0$ and for $p_1 = p_0 = 1/2$ X is a simple symmetric random walk.

Finding the marginal transition probabilities for the R-process requires more work. We find the marginal distribution of R by a direct calculation first, leaving an alternative proof via an intertwining argument until the next section.

First we prove the following lemma.

Lemma 3.15. *Consider $Z^{(q,\mu)} = (X_n, R_n; n \geq 0)$, $q \in (0,1]$ and $\mu \geq 0$, with $(X_0, R_0) = (0,0)$. For all admissible R-paths $\{r_1, ..., r_n\}$ (i.e. $r_m - r_{m-1} = \pm 1$ and $r_m \geq 0$ for $1 \leq m \leq n$) and $x \in \{-r_n, -r_n + 2, ..., r_n - 2, r_n\}$ the following is true*

$$\mathbb{P}(X_n = x, R_n = r_n, ..., R_1 = r_1) = p_1^{\frac{n+x}{2}} p_0^{\frac{n-x}{2}}. \tag{3.34}$$

In particular, when $p_1 = p_0 = 1/2$, the above probability is equal to $(1/2)^n$.

Proof. We prove the lemma by induction. It is trivially true for $n = 1$ as $\mathbb{P}(X_1 = 1, R_1 = 1 | X_0 = 0, R_0 = 0) = p_1$ and $\mathbb{P}(X_1 = -1, R_1 = 1 | X_0 = 0, R_0 = 0) = p_0$. Suppose now it is true for some $n > 1$ for any admissible choices of $\{r_1, ..., r_n\}$ and x_n. Then for $r = r_n \pm 1$

and x, s.t. $(x,r) \in \ell_r$, we write

$$\begin{aligned}
&\mathbb{P}(X_{n+1} = x, R_{n+1} = r, R_n = r_n, \ldots, R_1 = r_1) \\
&= \sum_{x_n \in \{x \pm 1\}} \mathbb{P}(X_{n+1} = x, R_{n+1} = r | X_n = x_n, R_n = r_n) \mathbb{P}(X_k = x_n, R_n = r_n, \ldots, R_1 = r_1) \ .
\end{aligned}$$

Using our assumption that (3.34) is true for n, the above is equal to

$$p_1^{\frac{n+x+1}{2}} p_0^{\frac{n-x-1}{2}} P((x+1, r_n), (x,r)) + p_1^{\frac{n+x-1}{2}} p_0^{\frac{n-x+1}{2}} P((x-1, r_n), (x,r)) \ . \tag{3.35}$$

Substituting corresponding transition probabilities, (3.31) or (3.32), we find that (3.35) is equal to $p_1^{(n+x+1)/2} p_0^{(n-x+1)/2}$, for $r = r_n \pm 1$. This proves the inductive step and, hence, the lemma. □

Proposition 3.16. *The marginal process $R = (R_n; n \geq 0)$ of the bivariate chain $Z^{(q,\mu)}$, for $q \in (0,1]$, is distributed as a dBES^3 started at 0 if $p_0 = p_1 = 1/2$ and as a dBES$^3(\mu)$ started at 0 if $p_1 - p_0 = \mu$.*

Proof. For any $n \geq 1$ we have

$$\mathbb{P}(R_{n+1} = r' | R_n = r, \ldots R_1 = r_1) = \frac{\mathbb{P}(R_{n+1} = r', R_n = r, \ldots R_1 = r_1)}{\mathbb{P}(R_n = r, \ldots, R_1 = r_1)} \ , \tag{3.36}$$

where $r' = r \pm 1$. But using Lemma 3.15, we can write for any $n \geq 1$

$$\begin{aligned}
\mathbb{P}(R_n = r_n, \ldots, R_1 = r_1) &= \sum_{x:(x,r_n) \in \ell_{r_n}} \mathbb{P}(X_n = x, R_n = r_n, \ldots, R_1 = r_1) \\
&= \sum_{x:(x,r_n) \in \ell_{r_n}} p_1^{\frac{n+x}{2}} p_0^{\frac{n-x}{2}} = \frac{1}{2^n}(r_n + 1) \quad \text{if } p_1 = p_2 \ , \\
&= p_1^{\frac{n-r_n}{2}} p_0^{\frac{n+r_n}{2}} \frac{(p_1/p_0)^{r_n+1} - 1}{p_1/p_0 - 1} \quad \text{if } p_1 \neq p_0 \ .
\end{aligned} \tag{3.37}$$

Substituting (3.37) in (3.36) and simplifying gives the required probabilities. □

Finally we prove a result concerning the behaviour of the dBES^3-process in general.

Proposition 3.17. *Let R be distributed as $dBES^3$ started at the origin. Then, for any $n \geq 1$ and $r \in \{n, n-2, ..., 1\}$ if n is odd, and $r \in \{n, n-2, ..., 0\}$ if n is even, we have*

$$\mathbb{P}(R_n = r) = \left(\begin{array}{c} n \\ \frac{n+r}{2} \end{array} \right) \frac{2(r+1)}{n+r+2} \frac{1}{2^n}(r+1)$$

if $\mu = 0$, and

$$\mathbb{P}(R_n = r) = \left(\begin{array}{c} n \\ \frac{n+r}{2} \end{array} \right) \frac{2(r+1)}{n+r+2} \frac{(p_1/p_0)^{r+1} - 1}{p_1/p_0 - 1} p_1^{\frac{n-r}{2}} p_0^{\frac{n+r}{2}}$$

if $\mu > 0$. In the above, as usual, we write $\left(\begin{array}{c} n \\ k \end{array} \right)$ for $\frac{n!}{(n-k)!k!}$.

Proof. The proof proceeds by induction. The above is trivially true for $n = r = 1$. For any $m \in \mathbb{N}$ and an admissible value of r, as described above, we note that

$$\begin{aligned} &\mathbb{P}(R_{m+1} = r) \\ &= \mathbb{P}(R_{m+1} = r | R_m = r-1)\mathbb{P}(R_m = r-1) + \mathbb{P}(R_{m+1} = r | R_m = r+1)\mathbb{P}(R_m = r+1) \, . \end{aligned}$$

Then, substituting expressions (3.17) and (3.18) for the transition probabilities of R in the above and assuming that the statement of the proposition is true for m, one proves via a direct calculation that it is also true for $m+1$. □

Note that if we compare equations in the statement of the proposition with equations (3.37), we conclude that for any $n \geq 1$

$$\#|\text{admissible R-paths } (r_1, ..., r_n = r)| = \left(\begin{array}{c} n \\ \frac{n+r}{2} \end{array} \right) \frac{2(r+1)}{n+r+2} \, .$$

3.6.2 Intertwining

In this section we show that the intertwining criteria of Rogers and Pitman can in fact be employed to show that the marginal R-process is Markovian and is distributed as the $\mathrm{dBES}^3(\mu)$ process. We start by outlining sufficient conditions that need to be satisfied by transition probabilities of a bivariate Markov process in $\mathbf{W}$ with one of the marginals given by the $\mathrm{RW}(\mu)$ in order for the second marginal to be distributed as the $\mathrm{dBES}^3(\mu)$ process, for $\mu \geq 0$.

Lemma 3.18. *Let $Z^{(\mu)} = (X_n, R_n; n \geq 0)$ be a discrete homogenous Markov chain with state space $\boldsymbol{W}$ (3.33) and such that X is distributed as a simple random walk with drift $p_1 - p_0 = \mu \geq 0$. Moreover, suppose that X is Markov with respect to the natural filtration of $Z^{(\mu)}$. Let $\lambda^\mu(r, \cdot)$, for any $r \in \mathbb{N}$, be a measure on $\boldsymbol{W}$ given by*

$$\lambda^0(r, (x', r')) := \lambda(r, (x', r')) = \frac{1}{r+1} \mathbf{1}_{\{r'=r\}} \tag{3.38}$$

if $\mu = 0$, and

$$\lambda^\mu(r, (x', r')) = p_1^{\frac{r+x'}{2}} p_0^{\frac{r-x'}{2}} \frac{p_1 - p_0}{p_1^{r+1} - p_0^{r+1}} \mathbf{1}_{\{r'=r\}} \tag{3.39}$$

if $\mu \geq 0$. Then for each $n \geq 1$ the n-step transition probabilities of Z and $dBES^3(\mu)$, denoted P_n^μ and Q_n^μ respectively, satisfy the following intertwining relationship

$$\sum_{(x'', r'') \in \boldsymbol{W}} \lambda^\mu(r', (x'', r'')) P_n^\mu((x'', r''), (x, r)) = \sum_{r'' \geq 0} Q_n^\mu(r', r'') \lambda^\mu(r'', (x, r)) \tag{3.40}$$

for all $n \geq 1$, if and only if the first-order transition probabilities of the two-dimensional process satisfy

$$\begin{aligned} &\frac{1}{p_0} P^\mu((x+1, r+1), (x, r)) + \frac{1}{p_1} P^\mu((x-1, r+1), (x, r)) = 1 \ , \\ &\frac{1}{p_1} P^\mu((x-1, r-1), (x, r)) + \frac{1}{p_0} P^\mu((x-1, r+1), (x, r)) = 1 \end{aligned} \tag{3.41}$$

for all $(x, r) \in \boldsymbol{W} \backslash \{(0, 0)\}$.

Proof. First note that (3.40) is equivalent to

$$\sum_{x':(x',r')\in\ell_{r'}} \lambda^{\mu}(x',r')P^{\mu}_{n}((x',r'),(x,r)) = Q^{\mu}_{n}(r',r)\lambda^{\mu}(x,r)\,, \tag{3.42}$$

where $\lambda(x,r) = \lambda(r,(x,r))$. Now, by using expressions (3.17) and (3.18) for Q^{μ}, substituting $r' = r \pm 1$ and simplifying, one finds that the above holds for $n = 1$ if and only if (3.41) is true. This proves the necessity. To prove sufficiency proceed by induction and suppose that (3.42) is true for all $1 \le n \le k$ for some $k \in \mathbb{N}$. Then, using the fact that, being transition probabilities of Markov processes, both $(P^{\mu}_{n}; n \ge 1)$ and $(Q^{\mu}_{n}; n \ge 1)$ satisfy Chapman-Kolmogorov equations, we write

$$\begin{aligned}
&\sum_{x':(x',r')\in\ell_{r'}} \lambda^{\mu}(x',r')P^{\mu}_{k+1}((x',r'),(x,r)) = \\
&\quad= \sum_{x':(x',r')\in\ell_{r'}} \lambda^{\mu}(x',r') \sum_{(x'',r'')\in\mathbf{W}} P^{\mu}_{k}((x',r'),(x'',r''))P^{\mu}((x'',r''),(x,r)) \\
&\quad= \sum_{(x'',r'')\in\mathbf{W}} \left[\sum_{x':(x',r')\in\ell_{r'}} \lambda^{\mu}(x',r')P^{\mu}_{k}((x',r'),(x'',r'')) \right] P^{\mu}((x'',r''),(x,r)) \\
&\quad= \sum_{(x'',r'')\in\mathbf{W}} Q^{\mu}_{k}(r',r'')\lambda^{\mu}(x'',r'')P^{\mu}((x'',r''),(x,r)) \\
&\quad= \sum_{r''\ge 0} Q^{\mu}_{k}(r',r'') \sum_{x'':(x'',r'')\in\ell_{r''}} \lambda^{\mu}(x'',r'')P^{\mu}((x'',r''),(x,r)) \\
&\quad= \sum_{r''\ge 0} Q^{\mu}_{k}(r',r'')Q^{\mu}(r'',r)\lambda^{\mu}(x,r) = Q^{\mu}_{k+1}(r',r)\lambda^{\mu}(x,r)\,,
\end{aligned}$$

which proves that (3.42) holds for $n = k + 1$ and, hence, by induction for all $n \in \mathbb{N}$. This concludes the inductive step and so the proof of our proposition. □

Proposition 3.19. *Let $Z^{(\mu)} = (X,R)$, for $\mu \ge 0$, be as in Lemma 3.18 and let $Z^{(\mu)}(0)$ be distributed according to $\lambda^{\mu}(r_0,\cdot)$, for some $r_0 \ge 0$. Suppose the intertwining condition* (3.40) *is satisfied. Then the marginal process R is distributed as $dBES^3(\mu)$ started at r_0.*

Proof. The proof essentially consists of proving a discrete version of the Pitman and Roger's Theorem 3.6. For $m_k > ... > m_2 > m_1 \ge 1$ and an admissible R-path $\{R_{m_1} = r_1, ..., R_{m_k} = r_k\}$, using expression (3.42), which follows from the assumption of intertwining, we have

$$
\begin{aligned}
&\mathbb{P}_{r_0}(R_{m_1} = r_1, ..., R_{m_k} = r_k) \\
&\quad = \sum_{x_k} ... \sum_{x_1} \sum_{x_0} \lambda(x_0, r_0) P_{m_1}((x_0, r_0), (x_1, r_1)) ... P_{m_k - m_{k-1}}((x_{k-1}, r_{k-1}), (x_k, r_k)) \\
&\quad = Q^{\mu}_{m_1}(r_0, r_1) \sum_{x_k} ... \sum_{x_1} \lambda(x_1, r_1) P_{m_2 - m_1}((x_1, r_1), (x_2, r_2)) ... P_{m_k - m_{k-1}}((x_{k-1}, r_{k-1}), (x_k, r_k)) \\
&\quad = ... \\
&\quad = Q^{\mu}_{m_1}(r_0, r_1) ... Q^{\mu}_{m_k - m_{k-1}}(r_{k-1}, r_k) \sum_{x_k} \lambda(r_k, x_k) \\
&\quad = Q^{\mu}_{m_1}(r_0, r_1) ... Q^{\mu}_{m_k - m_{k-1}}(r_{k-1}, r_k) \ ,
\end{aligned}
$$

which proves the proposition. □

Finally we state and prove

Proposition 3.20. *Let $Z^{(q,\mu)} = (X_n, R_n; n \geq 0)$, for $\mu \geq 0$, be a discrete homogeneous Markov chain with transition probabilities given by* (3.31), (3.32) *or* (3.20). *Then the transition probabilities of $Z^{(q,\mu)}$ and the dBES$^3(\mu)$-process are intertwined with respect to the Markov kernel λ given in Lemma 3.18. Moreover, if $Z^{(q,\mu)}$ is started according to the law $\lambda(r_0, \cdot)$, for some $r_0 \geq 0$, then the marginal process R is distributed as dBES$^3(\mu)$ started at r_0.*

Proof. One checks by a direct calculation that the transition probabilities of $Z^{(q,\mu)}$, for $q \in [0,1]$ and $\mu \geq 0$, satisfy (3.41). Then by Lemma 3.18 the transition probabilities of $Z^{(q,\mu)}$ and the dBES$^3(\mu)$-process satisfy the intertwining relationship (3.40). The last assertion of the proposition then follows by Proposition 3.19. □

We conclude this section with

Proposition 3.21. *Family of measures $(\nu^{\mu}_n; n \geq 1)$ on **W** given by*

$$
\nu^{\mu}_n(x, r) = Q^{\mu}_n(r_0, r) \lambda^{\mu}(x, r) \quad \textit{for any } r_0 \in \mathbb{N} \ ,
$$

where Q^{μ} is the transition function of dBES$^3(\mu)$ and $\lambda^{\mu}(x, r) = \lambda^{\mu}(r, (x, r))$ is as in

Lemma 3.18, forms an entrance law for the family of transition probabilities of the process $Z^{(q,\mu)}$ for $q \in [0,1]$ and $\mu \geq 0$, i.e. for all $n,m \geq 1$

$$\nu^{\mu}_{n+m}(x,r) = \sum_{(x',r')\in \boldsymbol{W}} \nu^{\mu}_{n}(x',r') P^{\mu}_{m}((x',r'),(x,r)), \quad (x,r) \in \boldsymbol{W}\,. \tag{3.43}$$

Proof. Start with the RHS of (3.43) and use the relationship (3.42)

$$\begin{aligned}
\sum_{(x',r')\in \mathbf{W}} \nu^{\mu}_{n}(x',r') P^{\mu}_{m}((x',r'),(x,r)) &= \sum_{r'\geq 1} Q^{\mu}_{n}(r_0,r') \sum_{x':(x',r')\in \ell_{r'}} \lambda^{\mu}(x',r') P^{\mu}_{m}((x',r'),(x,r)) \\
&= \sum_{r'\geq 1} Q^{\mu}(r_0,r') Q^{\mu}_{m}(r',r) \lambda^{\mu}(x,r) \\
&= Q^{\mu}_{n+m}(r_0,r) \lambda^{\mu}(x,r) = \nu_{n+m}(x,r)\,,
\end{aligned}$$

where we have used the fact that the family $(Q^{\mu}_{n}, n \geq 1)$ satisfies the Chapman-Kolmogorov equations.

□

3.6.3 Counterexample

In previous section we have established that for *all* $q \in [0,1]$ the transition probabilities of the bivariate process $Z^{(q,\mu)}$ and the dBES$^3(\mu)$-process are intertwined with respect to a certain Markov kernel and that consequently the R-marginal of $Z^{(q,\mu)}$ is Markov by the criteria of Pitman and Rogers. One also easily checks that for $q = 0$ and $q = 1$, in the symmetric set-up ($\mu = 0$), R is Markov through the Dynkin condition. A natural question arises of whether all the 2-dimensional Markov chains with statespace $\mathbf{W}$ and marginal processes (RW(μ), dBES$^3(\mu)$) satisfy either the Dynkin criterion or the intertwining condition? At least in the case $\mu = 0$ the answer is no, as the following proposition shows.

Proposition 3.22. *(counterexample) Let $Z = (Z_n; n \geq 0) = (X_n, R_n; n \geq 0)$, with $Z_0 = (0,0)$, be a discrete homogenous Markov chain taking values in* $\mathbf{W}$*. Suppose the bivariate* *transition probabilities satisfy*

$$(i)\quad P((r,r),(r+1,r+1)) = P((-r,r),(-r-1,r+1)) = \frac{1}{2}, \quad \forall r \geq 1\ ,$$

$$(ii)\quad P((x,r),(x+1,r+1)) = P((x,r),(x-1,r+1))$$
$$= P((x,r),(x+1,r-1)) = P((x,r),(x-1,r-1)) = \frac{1}{4}$$

for all $(x,r) \in W \backslash (\ell_2 \cup \{(0,0)\})$ *such that* $|x| \neq r$*, where* $\ell_2 = \{(2,2),(0,2),(-2,2)\}$,

$$(iii)\quad P((2,2),(1,1)) = P((-2,2),(-1,1)) \neq \frac{1}{4}, \frac{1}{3}\ ,$$

$$(iv)\quad P((2,2),(1,1)) + P((0,2),(1,1)) + P((0,2),(-1,1)) + P((-2,2),(-1,1)) = 1\ ,$$

$$(v)\quad P((2,2),(1,1)) + P((2,2),(1,3)) = P((-2,2),(-1,1)) + P((-2,2),(-1,3))$$
$$= P((0,2),(1,1)) + P((0,2),(1,3)) = \frac{1}{2}\ .$$

Then $X = (X_n; n \geq 1)$ *is distributed as the SSRW and* $R = (R_n; n \geq 0)$ *is distributed as a 3-d discrete Bessel process. Moreover, the two-dimensional chain* Z *and the marginal process* R *satisfy neither the Dynkin criterion nor the intertwining condition.*

Note that all that conditions (i)-(v) say is that all the jump probabilities of Z on the lattice **W** are equal to either $1/2$ or $1/4$, depending on whether the process is currently at the boundary or not, except for the jumps from the points $\ell_2 = \{(2,2),(0,2),(-2,2)\}$.

Proof. The statement about the distribution of X is an easy consequence of the properties $(i),(ii)$ and (v) of the transition probabilities.

We start by defining two families of probability measures on $\ell_r = \{(x,r) : x \in \{-r,-r+2,...,r-2,r\}\}$ for $r \geq 0$:

$$\mathscr{M}(r) = \{\mathscr{L}(X_m)|(R_m = r,...,R_1 = r_1);$$
$$(r_1,...,r_{m-1},r) \text{ runs over all admissible paths ending at } r,\ m \geq 1\}$$

and $\widetilde{\mathscr{M}(r)}$, defined as follows. If Y is any ℓ_r-valued random variable such that $\mathscr{L}(Y) \in$

$\widetilde{\mathscr{M}}(r)$, then, denoting $\mathbb{P}(Y = x) = \mu_x$ for $x \in \{-r, -r+2, ..., r-2, r\}$, we have $\mu_x + \mu_{-x} = \frac{2}{r+1}$ for $x \in \{r, r-2, ..., 2, 0\}$ if r is even and for $x \in \{r, r-2, ..., 1\}$ if r is odd.

Note that $\mathscr{M}(r)$ is just a collection of conditional laws of X_m, for all $m \geq 1$, given the path of R up to the time m and such that $R_m = r$. We will show that in fact $\mathscr{M}(r) = \widetilde{\mathscr{M}}(r)$, for all $r \geq 0$. We first show that if we start our two-dimensional process Z on ℓ_r, $r \geq 3$, according to any distribution in $\widetilde{\mathscr{M}}(r)$, then the leftward (and hence rightward) jump probability of the R-process is that of the dBES^3. For any $\mu \in \widetilde{\mathscr{M}}(r)$ we have

$$
\begin{aligned}
&\mathbb{P}(R_n = r-1 | R_{n-1} = r) \\
&\quad = \sum_{x:(x,r)\in\ell_r} [P((x,r),(x-1,r-1)) + P((x,r),(x+1,r-1))]\mu_x \\
&\quad\quad = \frac{1}{4}(\mu_r + \mu_{-r}) + \frac{1}{2}(\mu_{r-2} + \mu_{r-4} + ... + \mu_{-r+4} + \mu_{-r+2}) \\
&\quad\quad\quad = \frac{r}{2(r+1)} = Q(r, r-1)\,, \qquad (3.44)
\end{aligned}
$$

where we have used properties (i) and (ii) of the joint transition probabilities stated above and, as before, Q denotes transition probability function of the dBES^3-process.

Next we show that any $\mu \in \widetilde{\mathscr{M}}(r)$, for all $r \geq 3$, is mapped by P, the transition probability function of Z, to some $\nu \in \widetilde{\mathscr{M}}(r+1)$ and vice versa. By this we mean that the following holds

$$
\begin{aligned}
\mathbf{P}_{r,r+1}\mu &= \nu Q(r, r+1), \quad r \geq 3\,, \\
\mathbf{P}_{r+1,r}\nu' &= \mu' Q(r+1, r), \quad r \geq 2
\end{aligned}
$$

for some $\mu, \mu' \in \widetilde{\mathscr{M}}(r)$ and $\nu, \nu' \in \widetilde{\mathscr{M}}(r+1)$, where $\mathbf{P}_{r,r+1}$ is the $(r+1) \times (r+2)$ tridiagonal matrix of joint transition probabilities from points in ℓ_r to points in ℓ_{r+1} and $\mathbf{P}_{r+1,r}$ is the $(r+2) \times (r+1)$ tridiagonal matrix of joint transition probabilities between ℓ_{r+1} and ℓ_r. Alternatively, the two matrix equations above can be expressed as two systems of equations

$$
P((x+1,r),(x,r+1))\mu_{x+1} + P((x-1,r),(x,r+1))\mu_{x-1} = \nu_x Q(r, r+1) \qquad (3.45)
$$

for $x \in \{-r-1, -r+1, ..., r-1, r+1\}$ and

$$P((x+1, r+1), (x, r))\nu'_{x+1} + P((x-1, r+1), (x, r))\nu'_{x-1} = \mu'_x Q(r+1, r) \quad (3.46)$$

for $x \in \{-r, -r+2, ..., r-2, r\}$.

We show first that (3.45) holds. For any $\mu \in \widetilde{\mathscr{M}}(r)$, using (ii), one calculates (with the convention $\mu_x = 0$ for all $|x| > r+1$)

$$\nu_x = \frac{1}{4}(\mu_{x+1} + \mu_{x-1})\frac{1}{Q(r, r+1)} \quad \text{and} \quad \nu_{-x} = \frac{1}{4}(\mu_{-x-1} + \mu_{-x+1})\frac{1}{Q(r, r+1)} ,$$

thereby obtaining $\nu_x + \nu_{-x} = 2/(r+2)$ and so $\nu = (\nu_{r+1}, \nu_{r-1}, ..., \nu_{-r+1}, \nu_{-r-1}) \in \widetilde{\mathscr{M}}(r+1)$. Conversely, for any $\nu \in \widetilde{\mathscr{M}}(r+1)$ we have

$$\mu_k = \frac{1}{4}(\nu_{x+1} + \nu_{x-1})\frac{1}{Q(r+1, r)} \text{and} \qquad \mu_{-x} = \frac{1}{4}(\nu_{-x-1} + \nu_{-x+1})\frac{1}{Q(r+1, r)} . \quad (3.47)$$

It follows that $\mu_x + \mu_{-x} = 2/(r+1)$ and, consequently, $\mu = (\mu_r, \mu_{r-2}, ..., \mu_{-r+2}, \mu_{-r}) \in \widetilde{\mathscr{M}}(r)$ and (3.46) holds.

Combined with (3.44), this shows that, if Z is started at any ℓ_r, $r \geq 3$, according to any distribution μ in $\widetilde{\mathscr{M}}(r)$, then the marginal process R is Markov and is distributed as the dBES^3, at least until Z hits ℓ_2. Finally we need to check that we can start the chain at the origin and R will still be distributed as dBES^3.

First notice that $\mathscr{M}(1) = \widetilde{\mathscr{M}}(1)$ (since for any $\eta = (\eta_1, \eta_{-1}) \in \mathscr{M}(1)$ we have $\eta_1 + \eta_{-1} = 1 = 2/2$ and so $\nu \in \widetilde{\mathscr{M}}(1)$). Furthermore, any (conditional) probability distribution $(\eta_1, \eta_{-1}) \in \mathscr{M}(1)$ on the two points of ℓ_1 gives the right left and right R-jump probabilities because of condition (ii) above:

$$\left(\frac{1}{4} + \frac{1}{2}\right)(\eta_1 + \eta_{-1}) = \frac{3}{4} = Q(1, 2) .$$

Also any $(\eta_1, \eta_{-1}) \in \mathscr{M}(1)$ is mapped by P to some $(\mu_2, \mu_0, \mu_{-2}) \in \widetilde{\mathscr{M}}(2)$ (in the sense of (3.45) explained above), i.e.

$$\left[\frac{1}{4}\eta_1 + \frac{1}{4}\eta_{-1}\right]\frac{1}{Q(1, 2)} = \frac{1}{3} := \mu_0$$

and so $\mu_{-1}+\mu_{+1}=1-\mu_0=2/3$ and $(\mu_1,\mu_0,\mu_1)\in\widetilde{\mathscr{M}}(2)$. Because of conditions (iii) and (iv), for each such distribution on the points in ℓ_2 we have the correct marginal R distribution, i.e. for any $(\mu_2,1/3,2/3-\mu_2)=(\mu_2,\mu_0,\mu_{-2})\in\widetilde{\mathscr{M}}(2)$ we calculate

$$\begin{aligned}\mu_2 P((2,2),(1,1))+\mu_0[P((0,2),(1,1))+P((0,2),(-1,1))]+\mu_{-2}P((-2,2),(-1,1))\\ =\frac{2}{3}P((2,2),(1,1))+\frac{1}{3}(1-2P((2,2),(1,1)))=\frac{1}{3}=Q(2,1)\end{aligned}\tag{3.48}$$

Next we show that any $(\mu_2,1/3,2/3-\mu_2)\in\widetilde{\mathscr{M}}(2)$ is mapped by P to some $(\nu_3,\nu_1,\nu_{-1},\nu_{-3})\in\widetilde{\mathscr{M}}(3)$ and vice versa (again in the sense of (3.45) and (3.46)). This will show that in fact $\mathscr{M}(r)=\widetilde{\mathscr{M}}(r)$ for all $r\geq 1$, which is needed to complete our proof.

For any $(\mu_2,1/3,2/3-\mu_2)\in\widetilde{\mathscr{M}}(2)$ one has, using $P((2,2),(3,3))=P((-2,2),(-3,3))=1/2$ (which follows from (i))

$$\frac{\mu_2 P((2,2),(3,3))}{Q(2,3)}=\frac{3}{4}\mu_2:=\nu_3\ ,$$

$$\frac{1}{3}\frac{P((2,2),(1,3))+P((0,2),(1,3))}{Q(2,3)}=\frac{1}{2}[P((2,2),(1,3))+P((0,2),(1,3))]:=\nu_1\ ,$$

$$\begin{aligned}&\frac{1}{3}\frac{P((0,2),(-1,3))+P((-2,2),(-1,3))}{Q(2,3)}\\ &\qquad=\frac{1}{2}[P((0,2),(-1,3))+P((-2,2),(-1,3))]:=\nu_{-1}\ ,\end{aligned}$$

$$\frac{(2/3-\mu_2)P((-2,2),(-3,3))}{Q(2,3)}=\frac{3}{4}(2/3-\mu_2):=\nu_{-3}\ ,$$

where $(\nu_3,\nu_1,\nu_{-1},\nu_{-3})\in\mathscr{M}(3)$. It follows that $\nu_1+\nu_{-1}=1/2$ and, using (iv), that $\nu_3+\nu_{-3}=1/2$; hence, any $(\mu_2,1/3,2/3-\mu_2)\in\widetilde{\mathscr{M}}(2)$ is mapped by P to a point in $\widetilde{\mathscr{M}}(3)$. The converse is true by our previous calculation (3.47) with $r=2$.

To complete the proof we need to put all the pieces together. We have proved that for a chain Z started at the origin (or on any ℓ_r according to any $\mu\in\widetilde{\mathscr{M}}(r)$) we have $\mathscr{M}(r)=\widetilde{\mathscr{M}}(r)$ for all $r\geq 0$. That is, for any $r\geq 0$ the collection of conditional distributions of X_n, $n \geq 1$, given the R-path of the joint chain up to time n and ending at $R_n = r$ (i.e. all the possible conditional distributions on the points of ℓ_r) is equal to a certain family of probability distributions $\widetilde{\mathscr{M}}(r)$. It follows that

$$\begin{aligned}
&\mathbb{P}(R_{n+1} = r+1 | R_1 = r_1, ..., R_n = r) \\
&= \sum_{x:(x,r)\in \ell_r} \mathbb{P}(X_n = x | R_1 = r_1, ...R_n = r)\big[\mathbb{P}(X_{n+1} = x+1, R_{n+1} = r+1 | X_n = x, R_n = r) + \\
&\qquad \mathbb{P}(X_{n+1} = x-1, R_{n+1} = r+1 | X_n = x, R_n = r)\big] \\
&= \left(\frac{1}{4} + \frac{1}{2}\right)(\mu_r + \mu_{-r}) + \frac{1}{2}(\mu_{r-2} + \mu_{r-4} + ... + \mu_{-r+4} + \mu_{-r+2}) = \frac{r+2}{2(r+1)}
\end{aligned}$$

for some $\mu = (\mu_{-r}, \mu_{-r+2}, ..., \mu_{r-2}, \mu_r) \in \widetilde{\mathscr{M}}(r)$ and for all $r \geq 3$ and $r = 1$ and $n \geq 0$. Analogous result for $r = 2$ follows from calculation (3.48).

Finally notice that, on one hand, condition (iii) means that we cannot have $P((2,2),(1,3)) + P((2,0),(1,3)) = 1/2$ *and* $P((-2,2),(-1,3)) + P((2,0),(-1,3)) = 1/2$, i.e. the condition (3.41) of Lemma 3.15, necessary for the intertwining, cannot hold. At the other hand, (iii) means that $P((2,2),(1,1)) \neq 1/3 = Q(2,1)$ and so Dynkin condition fails also, as it requires $\sum_{x'} P((x,r),(x',r') = Q(r,r')$ for any x.

$\square$

3.7 Diffusion approximation

This section is devoted to finding a *diffusion approximation* for the discrete Markov chain $Z = Z^{(q,\mu)}$, for $q \in (0,1)$ and $\mu \geq 0$, we have discussed in the previous section. (Note that, as before, we will drop explicit dependence of Z on parameters q and μ to simplify our notation.) We will prove that, appropriately scaled, it converges to the unique solution of a martingale problem associated to a certain generator. However, discussion of the martingale problem itself is postponed until next chapter. We merely state the definition.

Definition 3.23. *(Solution to a martingale problem) We say that a probability measure* $\mathbb{P}$ *on continuous paths in the wedge* W (3.1) *is a solution to the martingale problem for*

$(\mathscr{G}, \nu_0)$ *if there exists a measurable* W*-valued process* $(Z_t; t \geq 0)$ *such that*

$$\nu_0(A) = \mathbb{P}(Z_0 \in A), \quad \forall A \in \mathscr{B}_W ,$$

where $\mathscr{B}_W$ *is the Borel* σ*-algebra of the subsets of* W*, and for all* f *in the domain* $\mathscr{D}_{\mathscr{G}}$ *of* $\mathscr{G}$ *the process*

$$f(Z_t) - f(Z_0) - \int_0^t \mathscr{G} f(Z_s) ds$$

is a $\mathbb{P}$*-martingale with respect to* $\mathscr{F}_t = \sigma(Z_s; s \leq t)$*, the natural filtration of* Z*. A martingale problem for* $(\mathscr{G}, \nu_0)$ *is well-posed if a solution exists and is unique.*

and

Theorem 3.24. *For all* $\theta \in [0, \infty)$ *and* $\mu \geq 0$ *the martingale problem for the generator*

$$\begin{aligned}\mathscr{G}^{(\theta,\mu)} = \frac{1}{2}\frac{\partial^2}{\partial x^2} + \frac{1}{2}\frac{\partial^2}{\partial r^2} + \left(\coth(\theta r) - e^{-x\theta}\operatorname{csch}(\theta r)\right)\frac{\partial^2}{\partial x \partial r} + \mu \frac{\partial}{\partial x} \\ + \left(\mu(\coth(\theta r) - e^{-x\theta}\operatorname{csch}(\theta r)) + \theta e^{-x\theta}\operatorname{csch}(\theta r)\right)\frac{\partial}{\partial r}\end{aligned} \tag{3.50}$$

started at any $z_0 \in W' = W \backslash \partial W$ *has a unique solution. Here* W' *denotes the interior or the wedge* W*, and the domain of the generator is taken to be the collection of all bounded, infinitely continuously differentiable functions on* W *with compact support, denoted by* $\mathscr{C}_c^\infty(W)$ *(see p. 115).*

Proof. See proofs of Propositions 4.3 and 4.7 in Chapter 4. □

We start by constructing a sequence of approximating continuous Markov chains $\{Z^{(n)}\}_{n\geq 1} = \{(X^{(n)}, R^{(n)})\}_{n \geq 1}$ as follows. For all $n \in \mathbb{N}$ let

$$Z_{k/n}^{(n)} = \frac{1}{\sqrt{n}} Z_k \quad \text{for} \quad k \in \mathbb{N} , \tag{3.51}$$

and define $Z_t^{(n)}$ for $t \in [\frac{k}{n}, \frac{k+1}{n}]$ by linear interpolation

$$Z_t^{(n)} = (nt-k) Z_{k/n}^{(n)} + (k+1-nt) Z_{(k+1)/n}^{(n)} = \frac{1}{\sqrt{n}} \left[(nt-k) Z_k + (k+1-nt) Z_{k+1} \right] . \tag{3.52}$$

Let $Z_0^{(n)} = \frac{1}{\sqrt{n}} Z_0$. The parameter $q \in [0,1)$ and the jump probabilities $p_0 = P(x, x-1)$ and $p_1 = P(x, x+1)$ are also scaled. For all $n \geq 1$ let

$$q = q^{(n)} = 1 - \frac{\theta}{\sqrt{n}}, \quad \theta \in [0, \infty)$$

and $p_1 = \frac{1}{2}(1 + \mu/\sqrt{n}) = 1 - p_0$ (so that $p_1 - p_0 = \frac{\mu}{\sqrt{n}}$). This way in the limit the parameter $q \in (0,1)$ gets mapped to $\theta \in (0,\infty)$; thus, $\theta = 0$ corresponds to the case of characterising the $\mathrm{BES}^3(\mu)$-process as the modulus of $\mathrm{BM}(\mu)$, while $\theta = \infty$ corresponds to the set-up of the Pitman's theorem.

So, for each $n \in \mathbb{N}$ $(Z_t^{(n)}; t \geq 0)$ is a Markov chain with piece-wise linear paths and values in a two-dimensional infinite wedge $W = \{(x,r) \in \mathbb{R} \times \mathbb{R}^+ : |x| \leq r\}$. Let $\mathbb{P}_{z_0}^n$ be the law of $Z^{(n)}$ started at z_0, and define by $P_t^n(\cdot,\cdot)$, $t > 0$, the transition probability functions defined with respect to $\mathbb{P}_{z_0}^n$. The aim of this section is to prove the following

Theorem 3.25. *The limit* $\lim_{n\to\infty} \mathbb{P}_0^n$ *exists and is the unique solution to the martingale problem associated to the infinitesimal generator* $\mathscr{G}^{(\theta,\mu)}$ *started at* $z_0 = 0$.

The proof of the above theorem requires a lemma.

Lemma 3.26. *For any* $z_0 \in W'$ *the limit* $\lim_{n\to\infty} \mathbb{P}_{z_0}^n$ *exists and is the unique solution to the* $(\mathscr{G}^{(\theta,\mu)}, \delta_{z_0})$*-martingale problem.*

In proving Lemma 3.26, and then Theorem 3.25, we take a classical route: we first prove that the infinitesimal coefficients of the generators of the scaled chains $Z^{(n)}$ converge in an appropriate sense to the coefficients of $\mathscr{G}^{(\theta,\mu)}$. Next we prove that the family of laws $\{\mathbb{P}_{z_0}^n\}_{n\geq 1}$, for any $z_0 \in W'$, associated to $\{Z^{(n)}\}_{n\geq 1}$ is pre-compact, which ensures that there exists a convergent subsequence $\{\mathbb{P}_{z_0}^{n_k}\}_{k\geq 1}$ and its limit is the solution to the $(\mathscr{G}^{(\theta,\mu)}, \delta_{z_0})$-martingale problem. We use the framework and results from Stroock and Varadhan [76].

Recall that a subset M of $\mathscr{P}(W)$, the set of all probability measures on subsets of W, is called *pre-compact* (or *locally compact*) if for each $\epsilon > 0$ there exists a compact set $K_\epsilon \subset W$ such that

$$\inf_{\mu \in M} \mu(K_\epsilon) \geq 1 - \epsilon \ .$$

We define the infinitesimal parameters of $Z^{(n)}$ as follows:

$$a^n_{z_i z_j}(z) = n \int_{|z-z'|<1} (z_i - z'_i)(z_j - z'_j) P^n_{1/n}(z, dz') \ ,$$

$$b^n_{z_i}(z) = n \int_{|z-z'|<1} (z_i - z'_i) P^n_{1/n}(z, dz') \ ,$$

where $z = (z_1, z_2) = (x, r)$, $z' = (z'_1, z'_2) = (x', y') \in W$ and $i, j \in \{1, 2\}$.

The following theorem combines Lemmas 11.2.1, 11.2.2 and Theorem 11.2.3 of Stroock and Varadhan [76].

Theorem 3.27. *Let Y be a discrete-time homogenous Markov chain on $\mathbb{R}^d$ and let $\{Y^{(n)}\}_{n\geq 1}$ be a sequence of continuous Markov processes constructed by scaling and interlacing analogously to* (3.51) *and* (3.52). *Let $(a^n_{ij})_{1\leq i,j\leq d}$ and $(b^n_i)_{1\leq i\leq d}$ be the infinitesimal variance and mean of $Y^{(n)}$ respectively. Suppose that the following conditions hold*

$$\lim_{n\to\infty} \sup_{|y|\leq R} \| a^n(y) - a(y) \| = 0 \ , \tag{3.53a}$$

$$\lim_{n\to\infty} \sup_{|y|\leq R} |b^n(y) - b(y)| = 0 \ , \tag{3.53b}$$

$$\sup_{n\geq 1} \sup_{y\in\mathbb{R}^d} (\| a^n(y) \| + |b^n(y)|) < \infty \tag{3.53c}$$

for all $R > 0$, where a is a real symmetric non-negative definite matrix with continuous entries and $b : \mathbb{R}^2 \to \mathbb{R}^d$ is a continuous function. For a square matrix A we define its operator norm as $\| A \| = \sup_{|v|^2=1} \| Av \| = \sup_{|v|^2=1} \langle Av, Av\rangle^{1/2}$. Finally suppose that

$$\lim_{n\to\infty} \sup_{y\in\mathbb{R}^d} \left(nP^n_{1/n}(y, \mathbb{R}^d \backslash B_y(\delta)) \right) = 0, \quad \textit{for all } \delta > 0 \tag{3.54}$$

holds, where $B_y(\delta)$ is a ball of radius $\delta > 0$ centred at y. Additionally, assume that the martingale problem for the generator with coefficients a and b and the domain $\mathscr{C}^\infty_c(\mathbb{R}^d)$ started at y_0 is well-posed and denote its unique solution by $\mathbb{P}_{y_0}$. Then $\mathbb{P}^n_{y_0}$ converges to $\mathbb{P}_{y_0}$ as $n \to \infty$ uniformly on compact subsets of $\mathbb{R}^d$. (By $\mathscr{C}^\infty_c(\mathbb{R}^d)$ we denote the collection of all infinitely continuously differentiable, bounded functions on $\mathbb{R}^d$ with compact support (see p. 115)).

Our aim now is to check that our family of processes $\{Z_t^{(n)}; t \geq 0\}_{n\geq 1}$, satisfies conditions (3.53)–(3.54) of the above theorem. This, combined with the fact that the martingale problem for $(\mathscr{G}^{(\theta,\mu)}, \delta_{z_0})$ is well-posed for all $z_0 \in W'$, will prove that the sequence of bivariate processes $\{Z^{(n)}\}_{n\geq 1}$ converges in law to the bivariate diffusion governed by the generator $\mathscr{G}^{(\theta,\mu)}$ and started at z_0. One notices that the results of Stroock and Varadhan apply to processes defined on the whole of $\mathbb{R}^d$, while our family of approximating Markov chains is defined on W and the supposed limiting diffusion is W'-valued. However, by Theorem 3.24 the martingale problem associated to the generator $\mathscr{G}^{(\theta,\mu)}$ acting on smooth functions with the support in the interior of the wedge is well-posed when started away from the boundary. Moreover, once started away from the boundary of the wedge, any such solution returns to the boundary with probability 0 (see Lemma 4.4 and Proposition 4.7). One verifies by studying the proofs of results in [76] that they are in fact applicable to our case.

Proof of Lemma 3.26. We start by calculating infinitesimal parameters of the chain $Z^{(n)}$, using transition probabilities (3.32). Noticing that

$$\mathbb{P}^n[Z^{(n)}_{(k+1)/n} = z \pm 1/\sqrt{n} | Z^{(n)}_{k/n} = z] = \mathbb{P}[Z_{k+1} = \sqrt{n}z \pm 1 | Z_k = \sqrt{n}z], \quad \forall k \geq 0 \,,$$

we calculate

$$\begin{aligned}
a^n_{xx}(x,r) = 1, \quad a^n_{rr}(x,r) &= 1, \quad b^n_x(x,r) = \mu \,, \\
a^n_{xr}(x,r) &= \frac{2p_1 q^{\sqrt{n}x+1} + 2p_0 q^{\sqrt{n}x-1} - q^{-(\sqrt{n}r+1)} - q^{\sqrt{n}r+1}}{q^{\sqrt{n}r+1} - q^{-(\sqrt{n}r+1)}} \,, \\
b^n_r(x,r) &= \mu \frac{q^{\sqrt{n}x+1} + q^{\sqrt{n}x-1} - q^{-(\sqrt{n}r+1)} - q^{\sqrt{n}r+1}}{q^{\sqrt{n}r+1} - q^{-(\sqrt{n}r+1)}} \\
&\quad + \sqrt{n} \frac{q^{\sqrt{n}x+1} - q^{\sqrt{n}x-1}}{q^{\sqrt{n}r+1} - q^{-(\sqrt{n}r+1)}} \,.
\end{aligned}$$

Recall that the coefficients of the generator of our target diffusion are given by

$$\begin{aligned}
a_{xx}(x,r) = 1, \quad a_{rr}(x,r) = 1, \quad b_x(x,r) = \mu \ , \\
a_{xr}(x,r) = \coth(\theta r) - e^{-x\theta}\operatorname{csch}(\theta r) = \frac{e^{\theta r} + e^{-\theta r} - 2e^{-\theta x}}{e^{\theta r} - e^{-\theta r}} \ , \\
\text{and} \quad b_r(x,r) = \mu\left(\coth(\theta r) - e^{-x\theta}\operatorname{csch}(\theta r)\right) + \theta e^{-x\theta}\operatorname{csch}(\theta r) \\
= \mu\frac{e^{\theta r} + e^{-\theta r} - 2e^{-\theta x}}{e^{\theta r} - e^{-\theta r}} + \frac{2\theta e^{-\theta x}}{e^{\theta r} - e^{-\theta r}} \ ,
\end{aligned}$$

where we have used exponential representation of hyperbolic functions. One notices that the drift coefficient of the R-process explodes to infinity at the origin. To gain control over it and a similarly unbounded, as $n \to \infty$, coefficient b_r^n, we consider a sequence of stopped processes $\{Z^{(n)}_{\tau^n_\epsilon}\}_{n\geq 1}$, where $Z^{(n)}_{\tau^n_\epsilon} = (Z^{(n)}_{t\wedge\tau^n_\epsilon}; t \geq 0)$ with $\tau^n_\epsilon = \inf\{t : R^{(n)}_t \leq \epsilon\}$.

For (3.53a) we have

$$\begin{aligned}
\| a^n(x,r) - a(x,r) \| &= \sup_{|v|^2=1} \left\| \begin{pmatrix} 0 & a^n_{xr}(x,r) - a_{xr}(x,r) \\ a^n_{xr}(x,r) - a_{xr}(x,r) & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \right\| \\
&= \sup_{|v|^2=1} [(a^n_{xr}(x,r) - a_{xr}(x,r))(v_1^2 + v_2^2)^{1/2}] = (a^n_{xr}(x,r) - a_{xr}(x,r))^2 \ .
\end{aligned}$$

And so for any $R > 0$ and $z = (x,r) \in W$

$$\lim_{n\to\infty} \sup_{|z|\leq R} \| a^n(x,r) - a(x,r) \| = \lim_{n\to\infty} \sup_{|z|\leq R} (a^n_{xr}(x,r) - a_{xr}(x,r)) \ .$$

Convergence above will follow if we show that $\lim_{n\to\infty} a^n_{xr}(z) = a_{xr}(z)$ uniformly on compact subsets of W. But this can be deduced by ubstituting $q^{(n)} = (1-\theta/\sqrt{n})$ into the expression for a^n_{xr} and using the fact that $\lim_{n\to\infty}(1 + x/n)^n = e^x$, for $x \in \mathbb{R}$, uniformly on compact subsets of $\mathbb{R}$. Similarly

$$\lim_{n\to\infty} \sup_{|z|\leq R} |b^n(x,r) - b(x,r)| = \lim_{n\to\infty} \sup_{|z|\leq R} |b^n_r(x,r) - b_r(x,r)| = 0 \ .$$

Finally for (3.53c) we have

$$\sup_n \sup_{z\in W_\epsilon} (\| a^n(z) \| + |b^n(z)|) \le \sup_n \sup_{z\in W_\epsilon} (\| a^n(z) \|) + \sup_n \sup_{z\in W_\epsilon} (|b^n(z)|) \ , \tag{3.55}$$

where $W_\epsilon = \{(x,r) \in W : r \ge \epsilon\}$. Let us study the two summands separately. For the first summand we calculate

$$\begin{aligned}&\sup_n \sup_{z\in W_\epsilon} (\| a^n(z) \|) \\ &\qquad = \sup_n \sup_{z\in W_\epsilon} \left[\sup_{|v|^2=1} (v_1^2 + v_2^2 + 4v_1v_2a^n_{xr}(z) + a^n_{xr}(z)^2(v_1^2+v_2^2))^{1/2} \right] \le 3^{1/2} \ ,\end{aligned}$$

where we have used the fact that for a vector $v \in \mathbb{R}$ such that $v_1^2+v_2^2 = 1$, $\sup_{|v|^2=1}\{v_1v_2\} = 1/4$ and the fact that $|a^n_{xr}(z)| < 1$ for all $n \ge 1$ and $z \in W$ (this can be verified by observing that $q^{\sqrt{n}x}$ is a decreasing function of x).
Consider now the second summand of (3.55). Note that expression $q^{-(\sqrt{n}r+1)} - q^{\sqrt{n}r+1}$, for $r > 0$, is an increasing function of r. Using this and $q^{-1} - q^1 = \frac{2\theta}{\sqrt{n}} + o(\frac{1}{n})$, we obtain for the second summand (and this is where stopping the process when the R^n-marginal becomes too small has an impact)

$$\begin{aligned}\sup_n \sup_{(x,r)\in W_\epsilon} (|b^n(x,r)|) &= \sup_{n\ge1} \sup_{z\in W_\epsilon} (b^n_r(x,r)) \\ &\le \sup_n \sup_{(x,r)\in W_\epsilon} \left(\sqrt{n} \frac{q^{\sqrt{n}x+1} - q^{\sqrt{n}x-1}}{q^{\sqrt{n}r+1} - q^{-(\sqrt{n}r+1)}} \right) + \mu \\ &\le \sup_n \left(\frac{q^{-\sqrt{n}\epsilon}(2\theta + o(1/\sqrt{n}))}{q^{-(\sqrt{n}\epsilon+1)} - q^{\sqrt{n}\epsilon+1}} \right) + \mu \\ &= \sup_n \left(\frac{2\theta + o(1/\sqrt{n})}{q^{-1} - q^{2\sqrt{n}\epsilon+2}} \right) + \mu \\ &\le C_{\theta,\epsilon} < \infty \ ,\end{aligned}$$

where $C_{\theta,\epsilon}$ is a finite constant depending on θ and ϵ but independent of n. The last two weak inequality signs follow from the fact that an error term associated to the convergence $\lim_{n\to\infty}(1+\frac{x}{n})^n = e^x$, for $x \in \mathbb{R}$, is of order $\frac{1}{n}$. We can now conclude that

$$\sup_n \sup_{(x,r)\in W_\epsilon} (\| a^n(x,r) \| + |b^n(x,r)|) < \infty$$

for all $\epsilon > 0$.

Finally we have to check that family of measures $(\mathbb{P}^n_{z_0}; n \geq 1)$ satisfies condition (3.54), which, together with a technical condition (3.53c), ensures that it is locally compact. We need to show that for all $\delta > 0$

$$\lim_{n\to\infty} \sup_{z\in W_\epsilon} \left(n P^n_{1/n}(z, W_\epsilon \backslash B_z(\delta)) \right) = 0 \ . \tag{3.56}$$

If $\delta > 1$, then $P^n_{1/n}(z, W_\epsilon \backslash B(z,\delta)) = 0$ for all $n \geq 1$ and any $z \in W_\epsilon$. If $\delta < 1$, then there exists $n_\delta \in \mathbb{N}$, more specifically $n_\delta = \min\{n \in \mathbb{N} : 1/n \leq \delta\}$, such that

$$P^n_{1/n}(z, W_\epsilon \backslash B_z(\delta)) = 1 - \sum_{z'\in B_z(\delta)} P^n_{1/n}(z, z') = 0 \ .$$

In other words, as n tends to infinity step size of $Z^{(n)}$ tends to zero and so eventually it becomes smaller than δ making probability in question 0. This establishes (3.56).

Hence, by Theorem 3.27

$$Z^{(n)}_{\tau^n_\epsilon} \xrightarrow{\mathscr{L}} Z_{\tau_\epsilon} \quad \text{as } n \to \infty \ ,$$

where $Z_{\tau_\epsilon} = (Z_{t\wedge\tau_\epsilon}; t \geq 0)$ with $\tau_\epsilon = \inf\{t : R_t \leq \epsilon\}$, is the solution to the $(\mathscr{G}^{(\theta,\mu)}, \delta_{z_0})$-martingale problem stopped when the R-process comes within distance ϵ of the origin. This is true for arbitrary small $\epsilon > 0$, and, as mentioned before, it will be proved in Chapter 4 that the solution to the martingale problem for $\mathscr{G}^{(\theta,\mu)}$ started away from the boundary of the wedge W revisits the boundary, and the origin in particular, with probability 0. Thus we must have

$$Z^{(n)} \xrightarrow{\mathscr{L}} Z \quad \text{as } n \to \infty \ ,$$

which proves Lemma 3.26. $\square$

In order to prove Theorem 3.25 we need the final lemma

Lemma 3.28. *Let $(\nu_t^{\mu,n}; t \in \frac{1}{n}\mathbb{N})$ be the entrance law of the process $Z^{(n)}$ started at the origin and let $(\nu_t^{\mu}; t > 0)$ be a family of probability measures on W with the densities given by*

$$\nu_t^{\mu}(x,r) = q_t^{\mu}(0,r)\lambda^{\mu}(x,r)\ ,$$

where q_t^{μ} is the transition density of the $BES^3(\mu)$ and λ^{μ} is given by

$$\lambda^0(r,x) = \frac{1}{2r}$$

for $\mu = 0$, and

$$\lambda^{\mu}(r,x) = \frac{\mu e^{\mu x}}{e^{\mu r} - e^{-\mu r}}$$

for $\mu > 0$. Then for any $t > 0$

$$\nu_t^{\mu,n} \Rightarrow \nu_t^{\mu} \quad \textit{as } n \to \infty.$$

Proof. We proved in Proposition 3.21 that the family of discrete measures on $\mathbf{W}$ $(\nu_m^{\mu}; m \geq 1)$ forming an entrance law for the process Z is given by $\nu_m^{\mu}(x,r) = Q_m^{\mu}(0,r)\lambda^{\mu}(r,x)$, where Q_m^{μ}, for $m \geq 1$, is the m-step transition function of dBES^3 and $\lambda^{\mu}(x,r) = \lambda^{\mu}(r,(x,r))$ is defined in Lemma 3.18. Then for all $n \geq 1$ the family of measures $(\nu_t^{\mu,n}; t \in \frac{1}{n}\mathbb{N})$, with

$$\nu_t^{\mu,n}(x,r) := Q_t^{\mu,n}(0,r)\lambda^{\mu,n}(x,r) := Q_{tn}^{\mu}(0,\sqrt{n}r)\lambda^{\mu}(\sqrt{n}x,\sqrt{n}r), \quad (x,r) \in \frac{1}{\sqrt{n}}\mathbf{W}, t \in \frac{1}{n}\mathbb{N}$$

forms an entrance law for $Z^{(n)}$, i.e. for each $t > 0$ $\nu_t^{\mu,n}$ is the law of $Z_t^{(n)} = (X_t^{(n)}, R_t^{(n)})$. Using Proposition 3.17 one obtains an explicit expression for $\nu_t^{\mu,n}$ for any $n \geq 1$, $\mu \geq 0$ and $t \in \frac{1}{\sqrt{n}}\mathbb{N}$. On the other hand the BES$^3(\mu)$ transition densities are well known:

$$q_t^{\mu}(0,r) = \frac{1}{\sqrt{2\pi t^3}}\frac{r}{\mu}\exp\left(-\frac{r^2}{2t}\right)\exp\left(-\frac{\mu^2 t}{2}\right)(\exp(\mu r) - \exp(-\mu r))$$

for $\mu > 0$, and

$$q_t(0,r) = \sqrt{\frac{2}{\pi t^3}} r^2 \exp\left(-\frac{r^2}{2t}\right)$$

for $\mu = 0$, and so we have an explicit expression for ν_t^μ also.

Now fix $t > 0$ and consider a sequence of random variables $\{(X_t^{(n)}, R_t^{(n)})\}_{n\geq 1} = \{Z_t^{(n)}\}_{n\geq 1}$. Let $(x_n, r_n) \in \frac{1}{\sqrt{n}}\mathbf{W}$, $n \geq 1$, be any sequence of points such that $(x_n, r_n) \to (x, r) \in W$ as $n \to \infty$, and $h_n = \frac{2}{\sqrt{n}}$ be a scaling parameter depending on the size of the lattice 'cells' forming the state space of the random variable $Z_t^{(n)}$ (recall that for a fixed $t \in \frac{1}{\sqrt{n}}\mathbb{N}$ $R_t^{(n)} = \frac{1}{\sqrt{n}} R_{nt} \in \{0, \frac{2}{\sqrt{n}}, ..., \frac{nt}{\sqrt{n}}\}$ if nt is even and $R_t^{(n)} \in \{1, \frac{3}{\sqrt{n}}, ..., \frac{nt}{\sqrt{n}}\}$ if nt is odd, and $X_t^{(n)} \in \{-R_t^{(n)}, -R_t^{(n)} + \frac{2}{\sqrt{n}}, ..., R_t^{(n)} - \frac{2}{\sqrt{n}}, R_t^{(n)}\}$). A result by Billingsley [13, Thm. 3.3, p.30] states that if

$$\frac{1}{h_n^2} Q_t^{\mu,n}(0, r_n) \lambda^{\mu,n}(x_n, r_n) \to q_t^\mu(0,r)\lambda^\mu(x,r) , \tag{3.57}$$

then the desired weak convergence of measures holds. Firstly we have

$$\frac{1}{h_n}\lambda^n(x,r) = \frac{\sqrt{n}}{2}\lambda(\sqrt{n}x, \sqrt{n}r) = \frac{\sqrt{n}}{2(\sqrt{n}r+1)} \to \frac{1}{2r}, \quad n \to \infty ,$$

and, recalling that $p_1 = \frac{1}{2}(1 + \mu/\sqrt{n}) = 1 - p_0$ and using $\lim_{m\to\infty}(1 + x/m)^m = e^x$, for $x \in \mathbb{R}$, we have

$$\begin{aligned}\frac{1}{h_n}\lambda^{\mu,n}(x,r) = \frac{\sqrt{n}}{2}\lambda^\mu(\sqrt{n}x, \sqrt{n}r) = p_1^{\frac{\sqrt{n}(r+x)}{2}} p_0^{\frac{\sqrt{n}(r-x)}{2}} \frac{p_1 - p_0}{p_1^{\sqrt{n}r+1} - p_0^{\sqrt{n}r+1}} \frac{\sqrt{n}}{2} \\ \to \frac{\mu e^{\mu x}}{e^{\mu r} - e^{-\mu r}}, \quad n \to \infty .\end{aligned}$$

At the other hand, using ([31, pp. 179-180])

$$\begin{pmatrix} 2n \\ n+x \end{pmatrix} \frac{1}{2^{2n}} \approx \frac{1}{\sqrt{\pi n}} e^{-\frac{x^2}{n}},$$

we obtain

$$\frac{1}{h_n} Q_{nt}(0, \sqrt{n}x) = \frac{1}{h_n} \frac{2}{2^{nt}} \frac{(\sqrt{n}r+1)^2}{nt+\sqrt{n}r+2} \begin{pmatrix} nt \\ \frac{nt+\sqrt{n}r}{2} \end{pmatrix} \approx \sqrt{\frac{2}{\pi t}} e^{-\frac{r^2}{2t}} \frac{(\sqrt{n}r+1)^2}{nt+\sqrt{n}r+2}$$

$$\to q_t(0,r), \quad n \to \infty \; .$$

Finally in a similar fashion we calculate

$$\frac{1}{h_n} Q^{\mu,n}_{nt}(0,r) = \frac{1}{h_n} \begin{pmatrix} nt \\ \frac{nt+\sqrt{n}r}{2} \end{pmatrix} \frac{2(\sqrt{n}r+1)}{nt+\sqrt{n}r+2} \frac{(p_1 p_0)^{\frac{nt}{2}} p_1^{-\frac{\sqrt{n}r}{2}} p_0^{\frac{\sqrt{n}r}{2}} (p_1^{\sqrt{n}r+1} - p_0^{\sqrt{n}r+1})}{p_0^{\sqrt{n}r}(p_1 - p_0)}$$

$$\approx \frac{1}{\sqrt{2\pi t}} e^{-\frac{r^2}{2t}} \frac{2(\sqrt{n}r+1)}{nt+\sqrt{n}r+2} \frac{\sqrt{n}}{\mu} \left(1 - \frac{\mu^2}{n}\right)^{\frac{nt}{2}} \frac{p_1^{-\frac{\sqrt{n}r}{2}} p_0^{\frac{\sqrt{n}r}{2}} (p_1^{\sqrt{n}r+1} - p_0^{\sqrt{n}r+1})}{p_0^{\sqrt{n}r}}$$

$$\to q^\mu_t(0,r), \quad n \to \infty \; ,$$

where we have used $(p_1 p_0)^{nt/2} = \frac{1}{2^{nt}} \left(1 - \frac{\mu^2}{n}\right)^{\frac{nt}{2}}$.

Putting all the ingredients together we obtain (3.57) and thus the statement of the lemma by the theorem of Billingsley. □

Finally we can prove Theorem 3.25.

Proof of Theorem 3.25. It is proved in Chapter 4 (see Corollary 4.16) that a family of measures $(\nu^\mu_t; t > 0)$ of Lemma 3.28 forms the unique entrance law for $Z^{(\theta,\mu)} = (X,R)$, the unique solution to the $(\mathscr{G}^{(\theta,\mu)}, \delta_{(0,0)})$-martingale problem. Thus, by Lemma 3.28 for each fixed $t > 0$ the distribution of the pair $Z^{(n)}_t = (X^{(n)}_t, R^{(n)}_t)$, viewed as a pair of random variables as opposed to a bivariate stochastic process, converges to the distribution of the pair $Z_t = (X_t, R_t)$ as n tends to infinity.

Thus, using Lemma 3.26, which states that, started at any $z_0 \in W'$, the coupling $Z^{(n)} = (X^{(n)}, R^{(n)})$ converges weakly to the unique solution of the $(\mathscr{G}^{(\theta,\mu)}, \delta_{z_0})$-martingale problem, completes the proof.

□

Remark. Using the theorems and techniques of this section, coupled with Lemma

3.28, we can easily prove that appropriately scaled dBES$^3(\mu)$-process converges to the continuous BES$^3(\mu)$.

Remark. Note that, by identifying the solution to the $(\mathscr{G}^{(\theta,\mu)}, \delta_{(0,0)})$-martingale problem as a diffusion limit of the discrete Markov chain $Z^{(q,\mu)} = (X, R)$, we automatically establish the fact that the marginal processes of the resulting diffusion are given by BM(μ) and BES$^3(\mu)$. In the next chapter we give a direct proof of this result.

Chapter 4

Pitman's theorem and radial part of the 3-d Brownian Motion II: bi-variate diffusions in a wedge

In this chapter we study a two-parameter family of bivariate diffusions with values in the infinite two-dimensional wedge $W = \{(x,r) : r \in \mathbb{R}^+, |x| \leq r\}$, which we have obtained as a scaling limit of certain Markov chains from representation theory in previous chapter. Let $Z^{(\theta,\mu)} = (X,R)$ be the process in question. For each $\theta \in [0,\infty)$ and $\mu \geq 0$ we aim to characterise $Z^{(\theta,\mu)} := Z^{(\theta)} := Z$ as a unique solution to a martingale problem associated to an infinitesimal generator given by

$$\mathscr{G}^{(\theta,\mu)} = \frac{1}{2}\frac{\partial^2}{\partial x^2} + \frac{1}{2}\frac{\partial^2}{\partial r^2} + \left(\coth(\theta r) - e^{-x\theta}\operatorname{csch}(\theta r)\right)\frac{\partial^2}{\partial x \partial r} + \mu\frac{\partial}{\partial x}$$
$$+ \left(\mu(\coth(\theta r) - e^{-x\theta}\operatorname{csch}(\theta r)) + \theta e^{-x\theta}\operatorname{csch}(\theta r)\right)\frac{\partial}{\partial r}\,, \quad (4.1)$$

and which we have already encountered at the end of the last chapter (see p. 100). Note that by letting $\theta \to 0$ we recover the generator (3.7). If, at the other hand, we let $\theta \to \infty$, we obtain the limiting generator $\hat{\mathscr{G}}^{(\infty)}_\mu = \frac{1}{2}\frac{\partial^2}{\partial x^2} + \frac{1}{2}\frac{\partial^2}{\partial r^2} + \frac{\partial^2}{\partial x \partial r} + \mu\frac{\partial}{\partial x} + \mu\frac{\partial}{\partial r}$ which (together with the boundary condition $\frac{\partial f}{\partial r}(-r,r) = 0$ for all f in the domain of $\hat{\mathscr{G}}^{(\infty)}_\mu$) describes the behaviour of the coupling $(\mathrm{BM}(\mu), \mathrm{BES}^3(\mu)) = (X,R) = (X, X-2J)$

with $J_t = \inf_{s \le t} X_t$, which is an equivalent representation of the Pitman's theorem; one should compare $\hat{\mathscr{G}}_\mu^{(\infty)}$ to the generator (3.8).

We then give a direct proof that, for all $\theta \in [0, \infty)$, the marginal distribution of the R-process is that of $\mathrm{BES}^3(\mu)$ and that, in particular in the symmetric case, $\mu = 0$, R is distributed as BES^3. To achieve this goal we show that for each $\theta \in [0, \infty)$ the semigroup associated to the process $Z^{(\theta,\mu)}$ and the semigroup of the $\mathrm{BES}^3(\mu)$-process are intertwined with respect to a certain kernel. The main result is the following

Theorem 4.1. *Let $(p_t^{(\theta,\mu)}; t > 0) := (p_t^\mu; t > 0)$ be the family of transition densities of the bivariate diffusion $Z^{(\theta,\mu)} = (R, X)$ associated to the infinitesimal generator given by* (4.1)*, and let $(q_t^\mu; t > 0)$ be the family of transition densities of the BES$^3(\mu)$ process. Then for all $\theta \in [0, \infty)$ and $\mu \ge 0$ the transition densities p_t^μ and q_t^μ satisfy the intertwining relationship*

$$\int_W \lambda^\mu(r, (x', r')) p_t^\mu((x', r'), (x'', r'')) dx' dr' = \int_0^\infty q_t^\mu(r, r') \lambda^\mu(r', (x'', r'')) dr' \,, \quad (4.2)$$

where λ^μ is a density of a Markov kernel Λ^μ from $(0, \infty)$ to W given by

$$\Lambda^0(r, A) := \Lambda(r, A) = \frac{1}{2r} Leb\{x : (x, r) \in A\} = \frac{1}{2r} \int_{-r}^{r} \mathbf{1}_{\{(x,r) \in A\}} \, dx \qquad (4.3)$$

for $\mu = 0$, and

$$\Lambda^\mu(r, A) = \int_A \frac{\mu e^{\mu x}}{e^{\mu r} - e^{-\mu r}} \mathbf{1}_{\{(x,r) \in A\}} \, dx \qquad (4.4)$$

for $\mu > 0$, for all $A \in \mathscr{B}_W$ and $r > 0$. Here $\mathscr{B}_W$ is the Borel sigma-algebra of the subsets of W.

In fact, the intertwining relationship (4.2) holds for $r = 0$ also, with the convention $\Lambda(0, \cdot) = \delta_o(\cdot)$, i.e. if $\Lambda(0, \cdot)$ is a measure on W with its mass concentrated on the origin (see Cor. 4.16).

Note that, if we denote the semigroups of $Z^{(\theta,\mu)}$ and $\mathrm{BES}^3(\mu)$ by P_t^μ and Q_t^μ, $t \ge 0$, respectively, then we can write (4.2) in operator notation as

$$\Lambda^\mu P_t^\mu = Q_t^\mu \Lambda^\mu$$

or

$$\int_W \Lambda^\mu(r,dz)P_t^\mu(z,A) = \int_0^\infty Q_t^\mu(r,dr')\Lambda^\mu(r',A)$$

for all $t>0, r>0$ and $A \in \mathscr{B}_W$.

Because of Theorem 3.6, an easy consequence of the above result is

Proposition 4.2. *Let the set-up be as in Theorem 4.1 and suppose that $Z^{(\theta,\mu)}$ is started at the origin, resp. according to the measure $\Lambda^\mu(r_0,\cdot)$, for some $r_0>0$. Then the marginal process $(R_t, t\geq 0)$ is distributed as the $BES^3(\mu)$-process started from the origin, resp. r_0. Moreover, for any $A\in\mathscr{B}_W$ we have*

$$\mathbb{P}^\mu(Z_t \in A | R_s, 0\leq s\leq t) = \int_A \lambda^\mu(R_t,z)dz = \Lambda^\mu(R_t,A)\ , \tag{4.5}$$

where Λ^μ is the intertwining kernel of Theorem 4.1.

In what follows we often omit superscripts indicating dependency on θ to simplify our notation. By default reader may assume that, unless otherwise stated, $\theta\in(0,\infty)$.

4.1 A martingale problem: starting away from the origin

For our discussion of the intertwining to be meaningful we first have to prove that there is a unique process $Z^{(\theta)}$ associated to the generator (4.1) for each value $\theta\in[0,\infty)$ and that, consequently, this process has an associated semigroup. This involves showing that the *martingale problem for $\mathscr{G}_\mu^{(\theta)}$* has a solution for all starting points and that, moreover, this solution is unique and therefore possesses the strong Markov property ([69, Thm. 21.1]). We treat the martingale problem started from zero and away from zero separately as the singularity at the origin needs a special approach. In fact, to prove the intertwining, it is enough to show that the martingale problem started *away from the origin* is well-posed. Once proved, the intertwining relationship (4.2), will allow us to find the entrance law for the process (X,R) started at the origin. However, to then show that this entrance law is unique, we need to prove uniqueness of the martingale problem started at the origin, which is, therefore, deferred till the end of the chapter.

The martingale approach to characterising Markov processes was introduced by Stroock and Varadhan [76]. Let Z be a time-homogenous Markov process with the statespace W, defined on some filtered probability space $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\geq 0}, \mathbb{P})$ and let $\mathscr{G}$ be its infinitesimal generator with the domain $\mathscr{D}_{\mathscr{G}}$. Then

$$\mathbb{E}_{Z_s}[f(Z_t)] = f(Z_s) + \mathbb{E}_{Z_s}\left[\int_s^t \mathscr{G}f(Z_s)ds\right] , \tag{4.6}$$

for all $t > s \geq 0$ and $f \in \mathscr{D}_{\mathscr{G}_\mu}$, where $\mathbb{E}_z$ is the expectation with respect to the law of Z started at z. Let

$$M_t^f = f(Z_t) - f(Z_0) - \int_0^t \mathscr{G}f(Z_u)du .$$

It follows that

$$\mathbb{E}[M_t^f|\mathscr{F}_s] = \mathbb{E}\left[f(Z_t) - f(Z_0) - \int_0^t \mathscr{G}f(Z_u)du\,\Big|\mathscr{F}_s\right]$$
$$= f(Z_s) - f(Z_0) - \int_0^s \mathscr{G}f(Z_u)du = M_s^f ,$$

i.e. M_t^f is an $\mathscr{F}_t$-martingale. Stroock and Varadhan reversed this argument: if for a given generator $\mathscr{G}_\mu$ there exists a probability measure $\mathbb{P}^\mu$ under which M_t^f is a martingale for a large enough class of functions, then there is a Markov process with generator $\mathscr{G}_\mu$ governed by $\mathbb{P}^\mu$. Measure $\mathbb{P}^\mu$ is called a *solution to a* $(\mathscr{G}, \nu_0)$*-martingale problem*, where ν_0 is the initial law of Z. If the solution to the $(\mathscr{G}^\mu, \nu_0)$ problem exists and is unique, then it is said to be *well-posed*. We say that a martingale problem for $\mathscr{G}_\mu$ is well-posed if a $(\mathscr{G}_\mu, \nu_0)$-martingale problem is well-posed for all $\nu_0 \in \mathscr{P}(W)$, where $\mathscr{P}(W)$ is the collection of all Borel measures on W. It is customary to call the process $(Z_t; t \geq 0)$ itself a solution to the $(\mathscr{G}_\mu, \nu_0)$-problem; we will use both definitions.

In what follows we write $\partial W = \{(r,x) : r \in \mathbb{R}^+, |x| = r\}$ for the boundary of the wedge W and $W' = W \backslash \partial W$ for its interior. We also write $a^\mu = (a_{ij}^\mu)_{i,j\in\{x,r\}}$ for the 2×2 matrix of the diffusion coefficients and $b^\mu = (b_x^\mu, b_r^\mu)$ for the vector of the drift coefficients of the generator (4.1). Whenever the dependency on μ is missing, one may assume that we are in the driftless case $\mu = 0$.

For $\mathbb{R}^d$ we denote by $\mathscr{C}_c^\infty(\mathbb{R}^d)$ a collection of all bounded infinitely continuously differentiable functions f on $\mathbb{R}^d$ with compact support. Define $\mathscr{C}_c^\infty(W)$ as follows. Each $f \in \mathscr{C}_c^\infty(W)$ is a bounded infinitely continuously differentiable function on W with compact support and such that there exists a number $\epsilon_f > 0$ such that $f(z) = 0$ and $\frac{\partial^k}{\partial x^k} f(z) = \frac{\partial^k}{\partial r^k} f(z) = 0$, for all $k \geq 1$ and $z \notin W_{\epsilon_f}$, where $W_{\epsilon_f} = \{(x,r) \in W : (r+x)/2 > \epsilon_f, (r-x)/2 > \epsilon_f\}$. In other words each f in $\mathscr{C}_c^\infty(W)$, together with its partial derivatives of all orders, vanishes in the vicinity of the boundary. In their original analysis Stroock and Varadhan considered processes defined on the whole of $\mathbb{R}^d$ with generators with domains given by $\mathscr{C}_c^\infty(\mathbb{R}^d)$ and which do not allow for the boundaries. However, in Lemma 4.4 and Proposition 4.7 below we prove that for $z_0 \in W'$ any solution to the $(\mathscr{G}_\mu, \delta_{z_0})$-martingale problem, for $\mu = 0$ and $\mu > 0$ respectively, does not hit the boundary ∂W with probability one. Moreover, we can continuously map the wedge W into the whole of $\mathbb{R}^2$, mapping the boundary ∂W to infinity, via a transformation involving log. Hence, the aforementioned results in [76] will still be applicable to our case. We take $\mathscr{D}_{\mathscr{G}_\mu^{(\theta)}} = \mathscr{D}_{\mathscr{G}} = \mathscr{C}_c^\infty(W)$ to be the domain of the generator $\mathscr{G}_\mu^{(\theta)}$ for all $\theta \in [0,\infty)$ and $\mu \geq 0$.

We consider the symmetric, $\mu = 0$, and the drifting, $\mu > 0$, cases separately, starting with the former.

4.1.1 Symmetric case, $\mu = 0$

For the symmetric case generator (4.1) takes the following form

$$\mathscr{G}^{(\theta)} = \frac{1}{2}\frac{\partial^2}{\partial x^2} + \frac{1}{2}\frac{\partial^2}{\partial r^2} + \Big(\coth(\theta r) - e^{-x\theta}\operatorname{csch}(\theta r)\Big)\frac{\partial^2}{\partial x \partial r} + \theta e^{-x\theta}\operatorname{csch}(\theta r)\frac{\partial}{\partial r}\,, \tag{4.7}$$

where we write $\mathscr{G}^{(\theta)} = \mathscr{G}$ for $\mathscr{G}_0^{(\theta)}$. Because of the singularity at the origin, we first consider a $\mathscr{G}$-martingale problem started away from the origin and prove

Proposition 4.3. *The martingale problem for $(\mathscr{G}^{(\theta)}, \delta_{z_0})$, $z_0 \in W'$, has a unique solution for all $\theta \in [0,\infty)$.*

We start our analysis with the following lemma concerning the behaviour of the process at the boundary.

Lemma 4.4. *Any solution to the martingale problem for* $(\mathscr{G}^{(\theta)}, \delta_{z_0})$*, for all* $z_0 \in W'$ *and* $\theta \in [0, \infty)$*, almost surely doesn't hit the boundary* $\partial W = \{(r, x) : r \in \mathbb{R}^+, |x| = r\}$.

Proof. In the case $\theta = 0$, corresponding to the characterisation of the BES^3-process as the radial part of the 3-d BM, the pair $(X_t, R_t; t \geq 0) = Z^{(0)}$ is related via

$$\begin{aligned} R_t &= \sqrt{X_t^2 + Y_t^2 + W_t^2} \\ \Longleftrightarrow \quad R_t^2 - X_t^2 &= Y^2 + W^2 \,, \end{aligned} \tag{4.8}$$

where (X_t, Y_t, W_t) are the coordinates of a standard 3-d Brownian motion. Note that the bi-variate process $Z_t^{(0)}$ hits the boundary at any time $t > 0$ if and only if $R_t^2 = X_t^2$, that is if and only if the LHS of the second equality above is zero. This only happens if the square root of the right hand side of (4.8), which is distributed as the radial part of the 2-d BM, or the BES^2-process, is zero. We know that for $d \geq 2$ BM^d, and so BES^d, issued from the origin never hits zero once it leaves it. Consequently we can conclude that the RHS of (4.8) never vanishes with probability 1 and so $Z^{(0)}$ never hits the boundary almost surely. Indeed one can check that $f(r, x) = \ln(r^2 - x^2)$ is harmonic with respect to the generator $\mathscr{G}^{(0)}$ (3.7) (this is not surprising once one remembers that $f(a) = \ln(a)$ is harmonic with respect to the generator of the 2-dimensional Bessel process (see for example [7, p. 90])).

Now that we have considered the extreme case $\theta = 0$, we deal with $\theta \in (0, \infty)$. We begin by rotating our coordinate system via a linear transformation

$$T : (r, x) \to \left(\frac{r - x}{2}, \frac{r + x}{2} \right) := (u, v) := y \tag{4.9}$$

and re-calculating the generator $\mathscr{G}$ in terms of the new variables $(u,v) \in [0,\infty)^2$

$$\mathscr{A}^{(\theta)} = \mathscr{A} = \frac{1-e^{-2\theta u}}{2(e^{2\theta v}-e^{-2\theta u})}\frac{\partial^2}{\partial u^2} + \frac{\theta}{e^{2\theta v}-e^{-2\theta u}}\frac{\partial}{\partial u} + \frac{e^{2\theta v}-1}{2(e^{2\theta v}-e^{-2\theta u})}\frac{\partial^2}{\partial v^2} + \frac{\theta}{e^{2\theta v}-e^{-2\theta u}}\frac{\partial}{\partial v} \,. \quad (4.10)$$

Note that the symmetry of the roles played by u and v is broken by the parameter θ. Indeed, as we let θ tend to 0, u and v in the generator become interchangeable.

Since the paths of the original and the transformed processes are in one-to-one correspondence, any solution to the $(\mathscr{G},\delta_{z_0})$-martingale problem, for each $z_0 = (r_0,x_0) \in W'$, corresponds to a solution to the $(\mathscr{A},\delta_{y_0})$-martingale problem, where $y_0 = (u_0,v_0)$ with $u_0 = (r_0-x_0)/2$ and $v_0 = (r_0+x_0)/2$. Let $Y = (Y_t; t \geq 0) = (U_t,V_t; t \geq 0)$ be any possible solution to the $(\mathscr{A},\delta_{y_0})$-martingale problem, where $y_0 = (u_0,v_0) \neq (0,0)$, and let $\mathbb{P}_{y_0}$ be the corresponding probability measure. We know that any solution to the $(\mathscr{A},\delta_{y_0})$-martingale problem corresponds to a weak solution to the corresponding system of stochastic differential equations started at y_0 (see, for example, [69, V.20]). Hence, $(Y_t; t \geq 0)$ must satisfy the following 2-dimensional SDE

$$dU_t = \sqrt{\frac{1-e^{-2\theta U_t}}{e^{2\theta V_t}-e^{-2\theta U_t}}}d\beta_t^1 + \frac{\theta}{e^{2\theta V_t}-e^{-2\theta U_t}}dt \,,$$
$$dV_t = \sqrt{\frac{e^{2\theta V_t}-1}{e^{2\theta V_t}-e^{-2\theta U_t}}}d\beta_t^2 + \frac{\theta}{e^{2\theta V_t}-e^{-2\theta U_t}}dt \,,$$

where $(\beta^1,\beta^2) = \beta$ is a standard 2-dimensional Brownian motion. Consider a function $g : [0,\infty)^2 \to \mathbb{R}$ given by

$$g(u,v) = \ln(1-e^{-2\theta u}) + \ln(e^{2\theta v}-1) \,.$$

By applying Itô's lemma to g, we see that $G_t := g(U_t,V_t)$ satisfies

$$G_t = \int_0^t F(U_s,V_s)d\beta_s, \quad t > 0 \,,$$

with

$$F(u,v) = \left(\frac{2\theta e^{-2\theta u}}{\sqrt{(1-e^{-2\theta u})(e^{2\theta v}-e^{-2\theta u})}}, \frac{2\theta e^{2\theta v}}{\sqrt{(e^{2\theta v}-1)(e^{2\theta v}-e^{-2\theta u})}} \right) ,$$

i.e. G_t is a local martingale under $\mathbb{P}_{y_0}$ with respect to the natural filtration of (U_t, V_t). Note that $U_t = 0$, resp. $V_t = 0$, if and only if $X_t = R_t$, resp. $X_t = -R_t$, i.e. when the original process hits the boundary. Moreover, $g(u,v) = -\infty$ if $u = 0$ or $v = 0$. However, we know that, being a local martingale, G_t cannot explode to plus or minus infinity in finite time. Hence, the pair (U,V) never hits the u- and v-axis, and so the original process (X,R) never hits the boundary ∂W of the wedge. Therefore, any solution to the $(\mathscr{G}, \delta_{z_0})$-martingale problem, for $z_0 \in W'$, lives in the interior of the wedge.

□

Let us now come back to our martingale problem. Usually, sufficiently regular, i.e. bounded and continuous, coefficients of the generator in question are enough to prove that a particular martingale problem has a solution and that it is unique. In view of this we see that the biggest obstacle in solving the $(\mathscr{G}, \delta_{z_0})$-martingale problem is the degeneracy of the R drift coefficient $b_r = \theta \operatorname{csch}(\theta r)$ which explodes to infinity at zero. However, a result of Stroock and Varadhan stated below gives us a way around this problem.

Theorem 4.5. *([76, Cor. 10.1.2]) Let $a : [0,\infty) \times \mathbb{R}^d \to a(t,x)$ be a symmetric non-negative definite $d \times d$ matrix and let $b : [0,\infty) \times \mathbb{R} \to b(t,x)$ be a vector in $\mathbb{R}^d$. Let a and b be locally bounded measurable functions and $\mathscr{G}$ the associated infinitesimal generator with $\mathscr{D}_{\mathscr{G}} = \mathscr{C}_c^\infty(\mathbb{R}^d)$. Assume that there exists an increasing sequence of bounded open sets $E_n \subset [0,\infty) \times \mathbb{R}^d$ with $\bigcup_n E_n = [0,\infty) \times \mathbb{R}^d$ and bounded measurable coefficients a^n and b^n such that $a^n \equiv a$ and $b^n \equiv b$ on E_n, and for all n the martingale problem for $\mathscr{G}_n$ (an infinitesimal generator with coefficients (a^n, b^n) and domain $\mathscr{C}_c^\infty(\mathbb{R}^d)$) starting from x is well-posed. Then for each x there is at most one solution to the martingale problem for a and b starting from x. Moreover if $\mathbb{P}_x^n$ is the solution for the martingale problem for $\mathscr{G}_n$ started at x and if $\tau_n = \inf\{t \geq 0 : (t, x(t)) \notin E_n\}$, then a solution for the martingale* *problem associated to* $(\mathscr{G}, \delta_x)$ *exists if and only if*

$$\lim_{n\to\infty} \mathbb{P}^n_x(\tau_n \le t) = 0 \quad \textit{for all } t \ge 0\ .$$

Finally, if the above holds, then the unique solution $\mathbb{P}_x$ *for* $(\mathscr{G}, \delta_x)$*-martingale problem is equal to* $\mathbb{P}^n_x$ *on* $\sigma(X_s; 0 \le s \le \tau_n)$ *for all* $n \ge 1$.

Theorem 4.5 allows us to *localise* our martingale problem by "dividing" W into open subsets W_n such that $\bigcup_n W_n = W'$ and considering, for all $n \ge 1$, a martingale problem associated to $\mathscr{G}$ "restricted" to W_n, a $\mathscr{G}_n$-martingale problem. If for each $n \ge 1$ the martingale problem for $(\mathscr{G}_n, \delta_{z_0})$, $z_0 \in W'$, is well-posed then the martingale problem for $\mathscr{G}$ started at the same point z_0 has at most one solution. Moreover a necessary condition is given to determine whether this solution exists at all.

To proceed we need another result by Stroock and Varadhan, which this time is concerned with martingale problems for generators with bounded coefficients.

Theorem 4.6. *([76, Thm. 7.2.1]) Let* $a : [0,\infty) \times \mathbb{R}^d \to a(t,x)$ *be a symmetric positive-definite* $d \times d$ *matrix and let* $b : [0,\infty) \times \mathbb{R} \to b(t,x)$ *be a vector in* $\mathbb{R}^d$*. Let* a *and* b *be bounded measurable functions and assume that for each* $t > 0$ *and* $x \in \mathbb{R}^d$

$$\inf_{0\le s\le t} \inf_{v\in\mathbb{R}^d\setminus\{0\}} \langle v, a(s,x)v\rangle/|v|^2 > 0 \tag{4.11a}$$

$$\lim_{y\to x} \sup_{0\le s\le t} \| a(s,y) - a(s,x) \| = 0\ , \tag{4.11b}$$

where for a square matrix A *we define* $\| A \| = \sup_{|v|^2=1} \| Av \|$.

Then the martingale problem for the generator with coefficients (a,b) *and domain* $\mathscr{C}^\infty_c(\mathbb{R}^d)$ *is well-posed.*

We can now prove Proposition 4.3.

Proof of Proposition 4.3. To prove the proposition we have to apply Theorems 4.5 and 4.6 to our case. For each $n \ge 1$ define

$$W_{1/n} := \{(x,r) \in W : (r+x)/2 > 1/n, (r-x)/2 > 1/n, r < 2n\}$$

and $\tau_n = \inf\{t : Z_t \notin W_{1/n}\}$. For all $n \geq 1$, $W_{1/n}$ is an open bounded (triangular) subset of W (the reason for the '2' in '$r < 2n$' is to ensure that $W_{1/n}$ is non-empty for $n = 1$), and τ_n is the first exit time of the process Z from it. Let a^n and b^n, for all $n \geq 1$, be functions on W defined as follows: $a^n(z) = a(z)$ and $b^n_r(z) = b_r(z)$, for $z \in W_{1/n}$, with a^n and b^n_r extending continuously, measurably, *boundedly* to the whole of W, such that $|a^n_{xr}(z)| \leq c_n < 1$ for some constant $c_n \in \mathbb{R}$. Then $a^n \equiv a$ and $b^n \equiv b$ on $W_{1/n}$ and $\bigcup_{n\geq 1} W_{1/n} = W'$. Denote by $\mathscr{G}_n$ a generator with coefficients (a^n, b^n). We check that such a generator satisfies properties (4.11a) and (4.11b) of Theorem 4.6 to show that the associated martingale problem started at $z_0 \in W'$ is well-posed. Note that both properties are invariant under the transformation involving *log* required to extend the generator to the whole of $\mathbb{R}^2$ in order to make the theorem applicable. For any $v = (v_1, v_2) \in \mathbb{R}^2$ and $z \in W'$ we have

$$\begin{aligned}\langle v, a^n(z)v\rangle \frac{1}{|v|^2} &= \langle (v_1, v_2), \begin{pmatrix} a^n_{xx}(z) & a^n_{xr}(z) \\ a^n_{xr}(z) & a^n_{rr}(z) \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \rangle \frac{1}{v_1^2 + v_2^2} \\ &\geq \frac{a^n_{xx}(z)v_1^2 - 2c_n|v_1 v_2| + a^n_{rr}(z)v_2^2}{v_1^2 + v_2^2} = \frac{(v_1 - v_2)^2}{v_1^2 + v_2^2} + \frac{2|v_1 v_2|(1-c)}{v_1^2 + v_2^2} > 0\ ,\end{aligned}$$

where we have used $a^n_{xx} = a^n_{rr} = 1$ and the fact that $|a^n_{xr}(z)| \leq c_n < 1$ for all $z \in W'$. The uniform ellipticity condition (4.11a) only fails at the boundary, which we know is not visited by the process, when it is started in the interior of the wedge.

Condition (4.11b) follows trivially from the fact that $a^n(z)$ is continuous in z and does not directly depend on t.

Finally, a^n and b^n are bounded measurable functions by construction. We conclude that all conditions of Theorem 4.6 are satisfied, and so the $(\mathscr{G}_n, \delta_{z_0})$-martingale problem is well-posed; for any $z_0 \in W'$ we denote its unique solution by $\mathbb{P}^n_{z_0}$.

By Lemma 4.4 the solution to the $(\mathscr{G}, \delta_{z_0})$, if it exists, hits the boundary, and in particular the origin, with probability zero. However, to ensure that $\lim_{n\to\infty} \tau_n = \infty$ we also need to check that the process (X,R) doesn't run off to infinity at the unbounded part of the wedge (i.e. that $X_t = \infty$ or $R_t = \infty$ for $t < \infty$). To show this, it is enough that the coefficients (a, b) of $\mathscr{G}$ satisfy the 'linear growth bounds' (see, for example, [76, Thm. 10.2.2])

$$\|a(z)\| \leq K(1+|z|^2)\,, \tag{4.12a}$$

$$\langle z, b(z)\rangle \leq K(1+|z|^2), \tag{4.12b}$$

for all for $z \in W'$, for some finite constant $K \in \mathbb{R}$. Inequality (4.12a) is trivially true, since all the diffusion coefficients of $\mathscr{G}$ are bounded. To see why (4.12b) is true, we consider the cases when $2\theta r < 1$ and when $2\theta r \geq 1$ separately. Starting with the former, note that by Taylor approximation, for $r \in (0, \frac{1}{2\theta})$, we have $e^{-2\theta r} < 1 - 2\theta r + \frac{(-2\theta r)^2}{2} < 1 - \theta r$, i.e. $1 - e^{-2\theta r} > \theta r$. Thus

$$\langle z, b(z)\rangle = r b_r(x,r) = \frac{2\theta r e^{-\theta(r+x)}}{1-e^{-2\theta r}} < \frac{2\theta r}{1-e^{-2\theta r}} < 2\,, \tag{4.13}$$

which clearly satisfies (4.12b) for any $z = (x,r) \in W'$ such that $2\theta r < 1$. On the other hand, for all $(x,r) \in W'$ such that $2\theta r \geq 1$ we have

$$\langle z, b(z)\rangle = \frac{2\theta r e^{-\theta(r+x)}}{1-e^{-2\theta r}} < \frac{2\theta r}{1-e^{-2\theta r}} < \frac{2\theta r}{1-e^{-1}}\,, \tag{4.14}$$

that is, $\langle z, b(z)\rangle$ is dominated by a linear function of r for large enough r. Of course, then (4.12b) is satisfied with some constant $K \in \mathbb{R}$ for all $z = (x,r) \in W'$ such that $2\theta r \geq 1$. Choosing a constant large enough to work for both (4.13) and (4.14), we obtain (4.12b).

We can now conclude that $\lim_{n\to\infty} \tau_n = \infty$ and so we have $\lim_{n\to\infty} \mathbb{P}^n_{z_0}(\tau_n \leq t) = 0$ for all $t \geq 0$.

Hence, all conditions of Theorem 4.5 are satisfied and so the martingale problem for our original generator $\mathscr{G}^{(\theta)}$ started at $z_0 \in W'$ has a unique solution; denote it by $\mathbb{P}_{z_0}$. Moreover, $\mathbb{P}_{z_0} = \mathbb{P}^n_{z_0}$ on $\mathscr{F}_{\tau_n} = \sigma(Z_t; 0 \leq t \leq \tau_n)$ and so $\lim_{n\to\infty} \mathbb{P}^n_{z_0} = \mathbb{P}_{z_0}$.

□

We have thus proved that the martingale problem for $\mathscr{G}^{(\theta,0)}$ is well-posed if we start away from the origin.

4.1.2 Drifting case, $\mu > 0$

In this section we consider the martingale problem associated to the generator $\mathscr{G}^{(\theta,\mu)} = \mathscr{G}^{\mu}$ (4.1) started away from the origin.

Proposition 4.7. *The martingale problem associated to $(\mathscr{G}^{\mu}, \delta_{z_0})$, $z_0 \in W'$, is well-posed and, moreover, the solution hits the domain boundary ∂W with probability zero.*

To prove the above proposition we need the following

Theorem 4.8. *([65, Thm. IX.1.10]) Let a be a field of symmetric and non-negative matrices, b and c fields of vectors such that a, b and $\langle c, ac\rangle$ are bounded. For any a and b we write $\mathscr{G}_{(a,b)}$ for the generator with coefficients a and b. There is a one-to-one and onto correspondence between the solutions to the $(\mathscr{G}_{(a,b)}, \delta_{x_0})$ and $(\mathscr{G}_{(a,b+ac)}, \delta_{x_0})$ martingale problems. If $\mathbb{P}$ and $\mathbb{Q}$ are the corresponding solutions and X is the associated process, then*

$$\frac{d\mathbb{Q}}{d\mathbb{P}}|_{\mathscr{F}_t} = \exp\left(\int_0^t c(X_s)d\bar{X}_s - \frac{1}{2}\int_0^t \langle c, ac\rangle(X_s)ds\right) ,$$

where $\bar{X}_s = X_s - \int_0^t b(X_s)ds$ and $\mathscr{F}_t = \sigma(X_s; 0 \le s \le t)$.

Proof of Proposition 4.7. We notice that, if we take $c = (\mu, 0)$, then the generators $\mathscr{G}$ (4.7) and $\mathscr{G}^{\mu}$ (4.1) are related exactly in the manner described in the theorem above, i.e. if (a, b) and (a^{μ}, b^{μ}) are the coefficients of $\mathscr{G}$ and $\mathscr{G}^{\mu}$ respectively, then clearly $a(z) = a^{\mu}(z)$ for all $z \in W$. At the other hand, one checks, using $a_{xx}(z) = 1$, that indeed

$$\begin{pmatrix} b^{\mu}_x(z) \\ b^{\mu}_r(z) \end{pmatrix} = \begin{pmatrix} 0 \\ b_r(z) \end{pmatrix} + \begin{pmatrix} a_{xx}(z) & a_{xr}(z) \\ a_{xr}(z) & a_{rr}(z) \end{pmatrix}\begin{pmatrix} \mu \\ 0 \end{pmatrix} = \begin{pmatrix} \mu \\ b_r(z) + \mu a_{xr}(z) \end{pmatrix} .$$

Now, for $n \ge 1$, let $\tau_n = \inf\{t : Z_t \notin W_{1/n}\}$, where, as before, $W_{1/n} = \{(x, r) \in W : (x + r)/2 > 1/n, (r - x)/2 > 1/n, r < 2n\}$, and define functions a^n and b^n on W like in the proof of Proposition 4.3: $a^n(z) = a(z)$ and $b^n_r(z) = b_r(z)$, for $z \in W_{1/n}$, with b^n_r and a_n extending continuously, measurably and boundedly to the whole of W', such that $|a^n_{xr}(z)| \le c_n < 1$ for some constant $c_n \in \mathbb{R}$. Let $\mathscr{G}_n$ and $\mathscr{G}^{\mu}_n$ be the generators with coefficients (a^n, b^n) and $(a^n, b^n + ca^n)$ respectively. From proof of Proposition 4.3 we know that the martingale problem for $\mathscr{G}_n$ started at $z_0 \in W' = W \backslash \partial W$ is well-posed; as before we denote the corresponding unique solution by $\mathbb{P}^n_{z_0}$. Then, since a^n, b^n and $\langle c, a^n c\rangle = \mu^2$ are bounded, by Theorem 4.8 the $(\mathscr{G}^\mu_n, \delta_{z_0})$- martingale problem is also well-posed; we denote its unique solution by $\mathbb{P}^{\mu,n}_{z_0}$. Again by Theorem 4.8 we have, for $t > 0$

$$\frac{d\mathbb{P}^{\mu,n}_{z_0}}{d\mathbb{P}^n_{z_0}}|_{\mathscr{F}_{t\wedge\tau_n}} = \exp\left(\mu X_{t\wedge\tau_n} - 1/2\mu^2 t\wedge\tau_n\right) ,$$

and, consequently,

$$\mathbb{P}^{\mu,n}_{z_0}(Z_{t\wedge\tau_n} \in dz) = \exp\left(\mu x - 1/2\mu^2 t\wedge\tau_n\right)\mathbb{P}^n_{z_0}(Z_{t\wedge\tau_n} \in dz) . \tag{4.15}$$

Notice how, modulo the stopping, the measures $\mathbb{P}^n_{z_0}$ and $\mathbb{P}^{\mu,n}_{z_0}$ are related in precisely the same way as the laws of the standard one-dimensional Brownian motion and the Brownian motion with drift μ.

Recall that $\mathbb{P}_{z_0}$, the unique solution to the $(\mathscr{G}, \delta_{z_0})$-martingale problem, satisfies $\mathbb{P}_{z_0}|_{\mathscr{F}_{\tau_n}} = \mathbb{P}^n_{z_0}|_{\mathscr{F}_{\tau_n}}$ for all $n \geq 1$, and $\lim_{n\to\infty}\mathbb{P}^n_{z_0} = \mathbb{P}_{z_0}$. Moreover, as discussed in the previous section, under $\mathbb{P}_{z_0}$ for all $z_0 \in W'$, the process Z doesn't hit the boundary or run off to infinity in finite time almost surely. In particular, $\lim_{n\to\infty}\mathbb{P}^n_{z_0}(\tau_n \leq t) = 0$ for all $t > 0$. Next, recall that under $\mathbb{P}_{z_0}$ the marginal process X is distributed as a standard Brownian motion; thus $\exp(\mu X_t - \frac{1}{2}\mu^2 t)$ is a positive local martingale (see, for example, [69, p. 55]) with a bounded expectation equal to $\exp(\mu X_0)$. So, denoting by $\partial W_{1/n}$ the boundary of the wedge $W_{1/n}$ and appealing to the reverse Fatou's lemma, we write

$$\begin{aligned}
\lim_{n\to\infty}\mathbb{P}^{\mu,n}_{z_0}(Z_{t\wedge\tau_n} \in \partial W_{1/n}) &= \lim_{n\to\infty}\mathbb{E}^n_{z_0}[\exp(\mu X_{t\wedge\tau_n} - 1/2\mu^2 t\wedge\tau_n)\,\mathbf{1}_{\{Z_{t\wedge\tau_n}\in\partial W_{1/n}\}}] \\
&= \lim_{n\to\infty}\mathbb{E}_{z_0}[\exp(\mu X_{t\wedge\tau_n} - 1/2\mu^2 t\wedge\tau_n)\,\mathbf{1}_{\{Z_{t\wedge\tau_n}\in\partial W_{1/n}\}}] \\
&\leq \mathbb{E}_{z_0}[\limsup_{n\to\infty}\exp(\mu X_{t\wedge\tau_n} - 1/2\mu^2 t\wedge\tau_n)\,\mathbf{1}_{\{Z_{t\wedge\tau_n}\in\partial W_{1/n}\}}] \\
&= \mathbb{E}_{z_0}[\exp(\mu X_t - 1/2\mu^2 t)\,\mathbf{1}_{\{Z_t\in\partial W \text{ or } Z_t=\infty\}}] = 0 ,
\end{aligned}$$

since under $\mathbb{P}$ we have $Z_t \notin \partial W$ and $Z_t < \infty$ a.s. This proves that under the measures $\mathbb{P}^{\mu,n}_{z_0}$, as $n\to\infty$, the process Z cannot escape all the triangular domains $W_{1/n}$ in finite time, and so we have $\lim_{n\to\infty} \mathbb{P}^{\mu,n}_{z_0}(\tau_n \le t) = 0$ for all $t > 0$. Then by Theorem 4.5 there exists a unique solution to the $(\mathscr{G}^{\mu}, \delta_{z_0})$-martingale problem, for all $z_0 \in W'$; denote it by $\mathbb{P}^{\mu}_{z_0}$. By the same theorem $\mathbb{P}^{\mu}_{z_0} = \mathbb{P}^{\mu,n}_{z_0}$ on $\mathscr{F}_{\tau_n}$ for all $n \ge 1$ and so $\lim_{n\to\infty} \mathbb{P}^{\mu,n}_{z_0} = \mathbb{P}^{\mu}_{z_0}$. It follows that

$$\frac{d\mathbb{P}^{\mu}_{z_0}}{d\mathbb{P}_{z_0}}|_{\mathscr{F}_t} = \exp\left(\mu X_t - 1/2\mu^2 t\right) \ ,$$

and, as a consequence the laws of Z under $\mathbb{P}^{\mu}_{z_0}$ and $\mathbb{P}_{z_0}$ are absolutely continuous for all $z_0 \in W'$. In particular, since under $\mathbb{P}_{z_0}$ the process Z hits the boundary or escapes to infinity with probability 0, then under $\mathbb{P}^{\mu}_{z_0}$ the process also stays away from the boundary and doesn't explode almost surely.

□

Remark. We have mentioned before that in case $\theta = \infty$, i.e. in the set-up of the Pitman's theorem, the process $Z^{(\infty,\mu)}$ hits the boundary of the wedge with probability one. It is therefore interesting to notice that $Z^{(\infty,\mu)}$ is the only process in the family of bi-variate diffusions $(Z^{(\theta,\mu)}, \theta \in [0,\infty))$, for any $\mu \ge 0$, that reaches the domain boundary with positive probability.

4.1.3 Martingale problem with any initial distribution

To conclude our discussion of the martingale problem we prove

Proposition 4.9. *The martingale problem for $(\mathscr{G}^{\mu}, \nu)$, where ν is any Borel measure on W', is well-posed.*

Proof. (Adapted from proof of [8, Thm. 2.1]) As before, let $\mathbb{P}^{\mu}_{z_0}$ be the unique solution to the $(\mathscr{G}^{\mu}, \delta_{r_0})$-martingale problem for any $z_0 \in W'$. Denote by $(\Omega, \{\mathscr{F}_t\}_{t\ge 0}, \mathscr{F})$ the underlying filtered probability space. For any Borel measure ν on W' there must be at least one solution to the $(\mathscr{G}^{\mu}, \nu)$-martingale problem; denote it by $\mathbb{P}^{\mu}_{\nu}$. Let $\mathbb{Q}^{\mu}_{\omega}$ for $\omega \in \mathscr{F}_0$ be the regular conditional probability of $\mathbb{P}^{\mu}_{\nu}$ given $Z(0,\omega) = Z_0(\omega)$. Then by [76, Thm. 6.1.3] (if we stop the process when it comes too close to the origin so that the coefficients of the generator are bounded) there exist a $\mathbb{P}^{\mu}_{\nu}$-null set $N \in \mathscr{F}_0$, such that $\mathbb{Q}^{\mu}_{\omega}$ is a solution to the $(\mathscr{G}_{\mu}, \delta_{Z_0(\omega)})$-martingale problem for all $\omega \notin N$. It follows that $\mathbb{Q}^\mu_\omega = \mathbb{P}^\mu_{Z_0(\omega)}$ and, in particular, we know that this solution is unique. But then

$$\mathbb{P}^\mu_\nu = \int_\Omega \mathbb{Q}^\mu_\omega d\nu(\omega) = \int_W \mathbb{P}^\mu_{z_0} d\nu(z_0) ,$$

and so $\mathbb{P}^\mu_\nu$ must be the unique solution to the $(\mathscr{G}_\mu, \nu)$-martingale problem. $\square$

4.2 Forward equation

In this section we study a forward equation associated to the generator $\mathscr{G}^\mu$. We show that the LHS and the RHS of (4.2) both solve the forward equation and, by proving that the solution is in fact unique, we prove the intertwining equality (4.2). We begin with some definitions.

Definition 4.10. *(Solution to forward equation) Let $\mathscr{P}(W)$ be a collection of Borel probability measures on W. A family of measures $(\nu_t; t > 0)$ in $\mathscr{P}(W)$ is a solution to the Kolmogorov forward equation (also called Fokker-Planck equation by physicists) for $(\mathscr{G}, \nu_0)$, $\nu_0 \in \mathscr{P}(W)$ if*

$$\nu_t f = \nu_0 f + \int_0^t \nu_s(\mathscr{G} f) ds, \qquad \forall f \in \mathscr{D}_\mathscr{G} , \tag{4.16}$$

where $\nu_t f = \nu_t f(z) = \int_{W'} f(z) d\nu_t(z)$.

Suppose that $\mathbb{P}$ is a solution to the martingale problem for $(\mathscr{G}, \nu_0)$ and Z_t is the associated process. Then

$$\mathbb{E}\left[f(Z_t) - \int_0^t \mathscr{G} f(X_s) ds \right] = \mathbb{E}[f(X_0)] ,$$

i.e. the family of one-dimensional distributions of Z $(\nu_t; t \geq 0)$ (that is $\nu_t(A) = \mathbb{P}(Z_t \in A)$, $\forall A \in \mathscr{B}_W$, $\forall t \geq 0$) solves the forward equation. The converse, however, does not need to be true, that is a solution to the forward equation need not solve the associated martingale problem.

Now, consider $\mathbb{P}^\mu$, the unique solution to the $(\mathscr{G}^\mu, \Lambda^\mu(r_0, \cdot))$-martingale problem for some $r_0 > 0$, where Λ^μ is the Markov kernel defined in Theorem 4.1. Then, in particular, for $\mu = 0$ the associated process Z is started on the line $\ell_{r_0} := \{(x,r) \in W : r = r_0\}$ according to the uniform distribution on $(-r_0, r_0)$. Then for each $A \in \mathscr{B}_W$ we can write

$$\nu_t(A) := \mathbb{P}^\mu(Z_t \in A) = \int_{W'} \Lambda^\mu(r_0, dz) P_t^\mu(z, A) = \Lambda^\mu P_t^\mu(r_0, A) \,.$$

It follows from the discussion above that $(\Lambda^\mu P_t^\mu(r_0,\cdot), t > 0)$ must solve the forward equation for $(\mathscr{G}^\mu, \Lambda^\mu(r_0,\cdot))$. Our strategy now is as follows. First we show that $(Q_t^\mu \Lambda^\mu(r_0,\cdot), t > 0)$ also solves the $(\mathscr{G}^\mu, \Lambda^\mu(r_0,\cdot))$-forward equation and then, by virtue of showing that the solution is unique, we prove equality (4.2).

Proposition 4.11. *For all* $r_0 > 0$ $(\Lambda^\mu Q_t^\mu(r_0,\cdot), t > 0)$ *solves the* $(\mathscr{G}^\mu, \Lambda^\mu(r_0,\cdot))$*-forward equation.*

Proof. We start by noticing that (4.16) can be re-written as $\frac{d}{dt}\nu_s f = \nu_s(\mathscr{G} f)$, or

$$\int f(z) \frac{d}{dt} \nu_s(dz) = \int \mathscr{G} f(z) \nu_s(dz) \,.$$

Recall that the *formal adjoint* of an operator $\mathscr{G}$ is an operator $\mathscr{G}^*$ such that

$$\int f(z)(\mathscr{G}^* g)(z) dz = \int g(z)(\mathscr{G} f)(z) dz \tag{4.17}$$

for all functions $f \in \mathscr{D}_{\mathscr{G}}$ and $g \in \mathscr{D}_{\mathscr{G}^*}$. We take $\mathscr{D}_{\mathscr{G}^*}$ to be the space of all bounded functions on W' twice continuously differentiable on the interior of the wedge and with continuous extensions of derivatives to the boundary. Therefore, for a solution $(\nu_t, t \geq 0)$ to the $\mathscr{G}$-forward equation, such that the density of ν_t belongs $\mathscr{D}_{\mathscr{G}^*}$, for all $t > 0$, we can write

$$\int_{W'} f(z) \frac{d}{dt} \nu_s(z) dz = \int_{W'} f(z)(\mathscr{G}^* \nu_s)(z) dz \tag{4.18}$$

for all $f \in \mathscr{D}_{\mathscr{G}}$. Now, let π_t^μ be the density of the kernel $Q_t^\mu \Lambda^\mu(r_0,\cdot)$, $r_0 > 0$, with respect to the Lebesgue measure, i.e. for any $A \in \mathscr{B}_W$

$$Q_t^\mu \Lambda^\mu(r_0, A) = \int_A \pi_t^\mu(r,x)\,dx\,dr \ .$$

Then, to verify that $Q^\mu \Lambda^\mu$ solves the forward equation, it suffices to check that

$$\frac{d}{dt}\pi_t^\mu = \mathscr{G}^* \pi_t^\mu \ . \tag{4.19}$$

First of all we can find π_t^μ explicitly; for all $A \in \mathscr{B}_W$ we have

$$Q_t^\mu \Lambda^\mu(r_0, A) = \int_A q_t^\mu(r_0, r)\lambda^\mu(x,r)\,dx\,dr = \int_A \pi_t^\mu(x,r)\,dx\,dr \ , \tag{4.20}$$

where q_t^μ is the transition density of $\mathrm{BES}^3(\mu)$, $\mu \geq 0$, and λ^μ is the density of the kernel Λ^μ; we write $\lambda^\mu(x,r)$ for $\lambda^\mu(r,(x,r))$, $(x,r) \in W$. Transition density q_t of BES^3 was given in Chapter 3 (eqn. (3.10)), while the transition density of $\mathrm{BES}^3(\mu)$ is also well-known and is given by

$$q_t^\mu(x,y) = e^{-\mu^2 t/2}\sqrt{\frac{x}{y}}\frac{I_{1/2}(\mu y)}{I_{1/2}(\mu x)} q_t(x,y) \ ,$$

where $I_{1/2}$ is the modified Bessel function of the first kind. Clearly, for any $r_0 > 0$ and $t > 0$ $\pi_t^\mu(\cdot) = q_t^\mu(r_0, \cdot)\lambda^\mu(\cdot)$, $\mu \geq 0$, is twice continuously-differentiable and one might check (by plotting a graph of the function or otherwise) that $\pi_t(x,r) \to 0$ as $x, r \to \infty$, for all $t > 0$. So π_t belongs to the domain of the adjoint $\mathscr{G}^*$, which we calculate next.

Lemma 4.12. *The adjoint operator of* $\mathscr{G}^\mu$ (4.4)*, for* $\mu \geq 0$*, is given by*

$$\begin{aligned}\mathscr{G}^{*\mu} = \frac{1}{2}\frac{\partial^2}{\partial x^2} + \frac{1}{2}\frac{\partial^2}{\partial r^2} + \frac{\partial^2}{\partial x \partial r}\Big(\coth(2\theta r) - e^{-2x\theta}\operatorname{csch}(2\theta r)\Big) - \mu\frac{\partial}{\partial x} \\ - \frac{\partial}{\partial r}\Big(\mu(\coth(2\theta r) - e^{-2x\theta}\operatorname{csch}(2\theta r)) + 2\theta e^{-2x\theta}\operatorname{csch}(2\theta r)\Big) \ .\end{aligned}$$

Proof. Consider equation-definition of the adjoint operator (4.17). The proof essentially consists of integrating the right-hand side of the identity twice and verifying that all the correction terms vanish due to the properties of the functions in the domain of

$\mathscr{G}^\mu$. In what follows we write ∂_x for $\partial/\partial x$, ∂^2_{xr} for $\partial^2/\partial x \partial r$ and so on. As before, we denote the coefficients of $\mathscr{G}^\mu$ by a^μ_{xr}, b^μ_r, b^μ_x, $a^\mu_{xx} = a^\mu_{rr} = 1/2$. Take any two functions f and g such that $f \in \mathscr{D}_{\mathscr{G}}$ and $g \in \mathscr{D}_{\mathscr{G}^*}$. Integrating by parts the RHS twice yields

$$\begin{aligned}
&\int_W g(z)(\mathscr{G}^\mu f)(z)dz = \int_W g(x,r)(\mathscr{G}^\mu f)(x,r)dxdr \\
&= \int_W f(x,r)(\mathscr{G}^{*\mu} g)(x,r)dxdr + \int_0^\infty \left[\frac{1}{2}\partial_x f(x,r)g(x,r) - \frac{1}{2}f(x,r)\partial_x g(x,r)\right]_{-r}^{r} dr \\
&\quad + \int_0^\infty \left[\frac{1}{2}\partial_r f(x,r)g(x,r) - \frac{1}{2}f(x,r)\partial_r g(x,r)\right]_{x}^{\infty} dx \\
&\quad + \int_{-\infty}^0 \left[\frac{1}{2}\partial_r f(x,r)g(x,r) - \frac{1}{2}f(x,r)\partial_r g(x,r)\right]_{-x}^{\infty} dx \\
&\quad + \int_0^\infty [g(x,r)a^\mu_{xr}(x,r)\partial_r f(x,r)]_{-r}^{r} dx \\
&\quad - \int_0^\infty [\partial_x(g(x,r)a^\mu_{xr}(x,r))f(x,r)]_x^\infty dr - \int_{-\infty}^0 [\partial_x(g(x,r)a^\mu_{xr}(x,r))f(x,r)]_{-x}^\infty dr \\
&\quad + \int_0^\infty [g(x,r)b^\mu_r(x,r)f(x,r)]_{-r}^{r} dr \\
&\quad + \int_0^\infty [g(x,r)b^\mu_r(x,r)f(x,r)]_x^\infty dr + \int_{-\infty}^0 [g(x,r)b^\mu_r(x,r)f(x,r)]_{-x}^\infty dr \,,
\end{aligned}$$

where we have used the fact that for any function h on W we have $\int_W h(z)dz = \int_0^\infty \int_{-r}^{r} h(x,r)dxdr = \int_0^\infty \int_x^\infty h(x,r)drdx + \int_{-\infty}^0 \int_{-x}^{-\infty} h(x,r)drdx$. But any function f in the domain of the generator $\mathscr{G}$, together with its first and second derivatives, vanish at infinity and, in particular, in the vicinity of the boundary ∂W. Hence all the correction terms above disappear, leaving us precisely with the LHS of (4.17).

$\square$

One can now verify by a direct, albeit tedious, calculation that (4.19) holds. Observing that $\partial_x a^\mu_{xr} = b^\mu_r - \mu a^\mu_{xr}$ and that $\partial_x \pi^\mu_t = \mu \pi^\mu_t$ (and in particular for $\mu = 0$

$\partial_x a_{xr} = b_r$ and $\partial_x \pi_t = 0$) we calculate the RHS

$$\frac{1}{2}\partial^2_{xx}\pi^\mu_t + \frac{1}{2}\partial^2_{rr}\pi^\mu_t + \partial^2_{xr}(a^\mu_{xr}\pi^\mu_t) - \mu\partial_x\pi^\mu_t - \partial_r(b^\mu_r\pi^\mu_t) = -\frac{1}{2}\mu^2\pi^\mu_t + \frac{1}{2}\partial^2_{rr}\pi^\mu_t$$
$$= \frac{1}{2}\lambda^\mu\partial^2_{rr}q^\mu_t - b_{bes_\mu}\lambda^\mu\partial_r q^\mu_t - \lambda^\mu(\mu^2 - b^2_{bes_\mu})q^\mu_t \ ,$$

where the second equality follows from the definition $\pi^\mu_t = \lambda^\mu q^\mu_t$ and easily verifiable identities $\partial_r\lambda^\mu = -b_{bes_\mu}\lambda^\mu$ and $\partial_r b_{bes_\mu} = \mu^2 - b^2_{bes_\mu}$, where b_{bes_μ} is the drift coefficient of an SDE corresponding to the $\mathrm{BES}^3(\mu)$-process (that is, $b_{bes_\mu}(r) = \mu\coth(\mu r)$ for $\mu > 0$ and $b_{bes}(r) = 1/r$ for $\mu = 0$), λ^μ, the density of the kernel Λ^μ, can be calculated from (4.4) and (4.3), and $(q^\mu_t; t \geq 0)$ is the transition density of the $\mathrm{BES}^3(\mu)$-process.

Next, recall that the transition density of the $\mathrm{BES}^3(\mu)$-process, for $\mu \geq 0$, must satisfy the forward equation for the appropriate generator, that is, one has $\frac{\partial}{\partial t}q^\mu_t = \frac{1}{2}\frac{\partial^2}{\partial r^2}q^\mu_t - \frac{\partial}{\partial r}(b_{bes_\mu}q^\mu_t)$. Using this, we calculate the LHS of (4.19)

$$\partial_t\pi^\mu_t = \partial_t(\lambda^\mu q^\mu_t) = \lambda^\mu\partial_t q^\mu_t = \frac{1}{2}\lambda^\mu\partial^2_{rr}q^\mu_t - b_{bes_\mu}\lambda^\mu\partial_r q^\mu_t - \lambda^\mu(\mu^2 - b^2_{bes_\mu})q^\mu_t \ ,$$

which proves the identity (4.19).

Moreover, by assumption $\mathbb{P}^\mu(Z_0 \in dz) = \Lambda^\mu(r_0, dz) = Q_0\Lambda^\mu(r_0, dz)$, which shows that $(Q^\mu_t\Lambda^\mu(r_0,\cdot), t > 0)$ solves the $(\mathscr{G}_\mu, \Lambda^\mu(r_0,\cdot))$-forward equation. Finally we need to prove that the $(\mathscr{G}_\mu, \nu_0)$-forward equation has in fact a unique solution. We need a result by Ethier and Kurtz [29].

□

Theorem 4.13. *([29, Prop. IV.9.19]) Let $\mathscr{C}_0(E)$ be a space of continuous functions vanishing at infinity, where E is locally compact and separable. Let $\mathscr{G}$ be a linear operator on $\mathscr{C}_0(E)$ satisfying the positive-maximum principle and such that the martingale problem for $\mathscr{G}$ is well-posed. Let $\mathscr{D}_\mathscr{G}$ be an algebra and dense in $\mathscr{C}_0(E)$. If μ^1_t and μ^2_t are two Borel probability measures on $(E, \mathscr{B}(E))$ satisfying forward equation for $\mathscr{G}$ and such that $\mu^1_0 = \mu^2_0$ then $\mu^1_t = \mu^2_t$ for all $t > 0$.*

In our case E is W', the interior of the wedge. Being an open subset of a locally compact separable Euclidian space $\mathbb{R}^2$, W' itself is locally compact and separable.

The boundary $\partial W = \{(x,r) : r \in \mathbb{R}^+, |x| = r\}$ and $\{r, x = \infty\}$ play the role of the point at infinity for W'. By $\mathscr{C}_0(W') := \mathscr{C}_0(W)$ we mean a collection of continuous functions on W such that any f in $\mathscr{C}_0(W)$ vanishes close to the boundary ∂W and also $\lim_{x,r\to\infty} f(x,r) = 0$. The domain $\mathscr{D}_{\mathscr{G}}$ of $\mathscr{G}$ is an algebra. By a 'linear operator on $\mathscr{C}_0(W)$' we mean a linear operator whose domain and range are subsets of $\mathscr{C}_0(W)$. Since the domain of $\mathscr{G}$ is the space of all bounded infinitely continuously differentiable functions with compact support, clearly $\mathscr{G}^\mu$ is a linear operator on $\mathscr{C}_0(W)$. It is now left to show that $\mathscr{C}_c^\infty(W)$ is dense in $\mathscr{C}_0(W)$ and that $\mathscr{G}^\mu$ satisfies the positive-maximum principle. We present this as two lemmas.

Lemma 4.14. *The generator $\mathscr{G}^\mu$ satisfies the positive-maximum principle, i.e. for any $f \in \mathscr{D}_{\mathscr{G}^\mu}$ if $z' \in W'$ is such that $f(z') = \sup_{z\in W'} f(z)$ and $f(z') \geq 0$, then*

$$\mathscr{G}^\mu f(z') \leq 0 \,. \tag{4.21}$$

Proof. We first observe that if $f \in \mathscr{D}_{\mathscr{G}^\mu}$ and $z' \in W'$ is such that $f(z') = \sup_{z\in W'} f(z)$ then, because f and all its partial derivatives vanish in the vicinity of the boundary, we automatically have $f(z') \geq 0$. For such z' the Hessian matrix

$$H = \begin{pmatrix} \partial^2_{xx} f(z') & \partial^2_{xr} f(z') \\ \partial^2_{xr} f(z') & \partial^2_{rr} f(z') \end{pmatrix}$$

is non positive-definite, i.e. $v^T H v \leq 0$ for all $v \in \mathbb{R}^2\backslash\{0\}$. Thus

$$\frac{1}{2}\partial^2_{xx} f(z') + a_{xr}(z')\partial^2_{xr} f(z') + \frac{1}{2}\partial^2_{rr} \leq 0 \,.$$

Also we must have $\partial_x f(z') = \partial_r f(z') = 0$ and so $\mathscr{G}_\mu f(z') \leq 0$, as required. $\square$

Lemma 4.15. *The space $\mathscr{C}_c^\infty(W)$ is dense in $\mathscr{C}_0(W)$, the space of all continuous functions on W' vanishing at infinity.*

Proof. Let $\mathscr{C}_c(W)$ be the space of all continuous functions on W with compact support. Then $\mathscr{C}_c^\infty(W) \subset \mathscr{C}_c(W) \subset \mathscr{C}_0(W)$. We first prove that $\mathscr{C}_c^\infty(W)$ is dense in $\mathscr{C}_c(W)$

and then prove that $\mathscr{C}_c(W)$ is dense in $\mathscr{C}_0(W)$. Then, by transitivity of denseness, we conclude that $\mathscr{C}_c^\infty(W)$ is dense in $\mathscr{C}_0(W)$.

We first note that $\mathscr{C}_c^\infty(W)$ separates the points of W', i.e. for any $z, z' \in W'$, such that $z \neq z'$, there exists a function f in $\mathscr{C}_c^\infty(W)$ such that $f(z) \neq f(z')$. To see why, take two points $z = (x, r), z' = (x', r') \in W'$, $z \neq z'$, such that $z, z' \in K$, where K is some subset of W. Without loss of generality we assume that $x \neq x'$. Let $f \in \mathscr{C}_c^\infty(W)$ be a function such that $f(z) = f(x, r) = x$ for all $z \in K$. Then $f(z) \neq f(z')$. Evidently, for any compact $K \subset W'$, there exists $f \in \mathscr{C}_c^\infty(W)$ such that $f = 1$ on K. Then by the locally compact version of the Stone-Weierstraß theorem (see, for example, [55, Lemma 2, p. 343]) $\mathscr{C}_c^\infty(W)$ is dense in $\mathscr{C}_c(W)$.

Now take $f \in \mathscr{C}_0(W)$. Since f is continuous and vanishes at infinity, for each $\epsilon > 0$ there is a subset K of W' such that $|f(z)| < \epsilon/2$ for all $z \notin K$. Let U be a compact subset of W' such that $K \subset U$ and define a function $g \in \mathscr{C}_c(W)$ with support U such that $g = f$ on K and such that $|g(z)| < \epsilon/2$ for $z \notin K$. Then $\|f(z) - g(z)\| \leq \epsilon$ for all $z \in W$, where $\|\cdot\|$ is the sup-norm. In other words, any $f \in \mathscr{C}_0(W)$ can be approximated by a function in $\mathscr{C}_c(W)$ and so $\mathscr{C}_c(W)$ is dense in $\in \mathscr{C}_0(W)$. By transitivity it follows that $\mathscr{C}_c^\infty(W)$ is dense in $\mathscr{C}_0(W)$.

$\square$

4.3 Intertwining

We can now finally prove Theorem 4.1.

Proof of Theorem 4.1. We have proved that for all $r_0 > 0$ the families of measures $(\Lambda^\mu P_t^\mu(r_0, \cdot), t \geq 0)$ and $(Q_t^\mu \Lambda^\mu(r_0, \cdot); t \geq 0)$ solve the $(\mathscr{G}^\mu, \Lambda^\mu(r_0, \cdot))$-forward equation. Moreover, trivially $\Lambda^\mu P_0^\mu = Q_0^\mu \Lambda^\mu$. It now follows from Theorem 4.13 and Lemmas 4.14 and 4.15 that

$$\Lambda^\mu P_t^\mu = Q_t^\mu \Lambda^\mu, \quad t > 0 .$$

$\square$

We can now find an entrance law for the process Z^μ started according to $\Lambda^\mu(r_0, \cdot)$, for $r_0 > 0$, and also for the process started at the origin. Because of the uniqueness of solution to the $(\mathscr{G}^\mu, \delta_{z_0})$-martingale problem for $z_0 \in W'$, this entrance law is unique for all $r_0 > 0$. We will show in the next section that there is only one way to start the process from the origin.

Corrolary 4.16. *For any $r_0 \geq 0$, the family of probability measures with densities $(\pi^\mu_t; t > 0)$ given by*

$$\pi_t(x,r) = q^\mu_t(r_0, r)\lambda^\mu(x,r) \ ,$$

forms an entrance law for the family of transition probability densities $(p^\mu_t; t > 0)$ satisfying the intertwining relationship (4.2)*, i.e. for all $t, s > 0$*

$$\pi^\mu_{t+s}(z) = \int_W \pi^\mu_t(z')p_s(z',z)dz' \ .$$

Proof. We first notice that the identity

$$\int_{-r'}^{r'} \lambda^\mu(r',x')p^\mu_t((r',x'),(r,x))dx' = q^\mu_t(r',r)\lambda^\mu(r,x) \ ,$$

where again we write $\lambda^\mu(r,x) = \lambda^\mu(x,(x,r))$, is an easy consequence of the intertwining (4.2) and the definition of λ^μ (4.3), (4.4). Using this, we have for all $r_0 \geq 0$

$$\begin{aligned}
\int_W \pi^\mu_t(x',r')p^\mu_s((x',r'),(x,r))dx'dr' \\
&= \int_0^\infty \int_{-r}^{+r} q^\mu_t(r_0,r')\lambda^\mu(x',r')p^\mu_s((x',r'),(x,r))dx'dr' \\
&= \int_0^\infty q^\mu_t(r_0,r')q^\mu_s(r',r)\lambda^\mu(r,x)dr' = q^\mu_{t+s}(r_0,r)\lambda^\mu(r,x) \\
&= \pi^\mu_{t+s}(x,r) \ .
\end{aligned}$$

□

Proof of Proposition 4.2. First of all we notice that from Corollary 4.16 it follows that

$$\int_A p^\mu_t(0,z)dz = \int_A q^\mu_t(0,r)\lambda^\mu(x,r)dxdr,$$

for all $A \in \mathscr{B}_W$ and $t > 0$. That is, the intertwining relationship (4.2) holds even for $r_0 = 0$ with the convention $\lambda^\mu(0,\cdot) = \delta_o(\cdot)$ (i.e. $\Lambda^\mu(0,\cdot)$ is a point mass concentrated on the origin).

One then easily checks other conditions of Theorem 3.6. Let $\phi : W \to (0,\infty)$ be defined as $\phi(x,r) = r$. Then for each bounded function f on $\mathbb{R}^+$ and $r \geq 0$ we have

$$\int_W \lambda^\mu(r,(x',r'))f(\phi(x',r'))dx'dr' = \int_{-r}^{r} \lambda^\mu(r,(x',r))f(r)dx' = f(r)\,,$$

that is for each bounded f defined on $\mathbb{R}^+$ we have $\Lambda^\mu \Phi f = f$, where $\Phi f = f \cdot \phi$. So, all conditions of Theorem 3.6 of Pitman and Rogers are satisfied and it follows that the marginal process R of $Z^{(\theta,\mu)} = (X,R)$ is a Markov process with transition densities $(q_t^\mu, t \geq 0)$ and started at $r_0 \geq 0$. In particular for $\mu = 0$ R is distributed as the 3-dimensional Bessel process and for $\mu > 0$ R is distributed as the $\text{BES}^3(\mu)$-process started at $r_0 \geq 0$. The filtering relationship (4.5) is also a direct consequence of Theorem 3.6.

$\square$

4.4 A martingale problem: starting at the origin

Finally, we address the question of the uniqueness of the process started from the origin. Having found an entrance law, we know that there is at least one such process; however, proving that this process is unique requires more work.

Symmetric case, $\mu = 0$. We first consider the martingale problem for the generator $\mathscr{G}^{(\theta,0)} = \mathscr{G}$ (4.7) started at the origin, which we denote by o. We prove

Proposition 4.17. *The martingale problem for $(\mathscr{G}^{(\theta,0)}, \delta_o)$ is well-posed.*

This situation is not as straightforward as the one discussed in Section 4.1 as the origin belongs to the statespace boundary.

Proof. Consider again the coordinate transformation (4.9). The look of the generator $\mathscr{A}$ (4.10) suggests that a time change might be appropriate. Recall that a *time change* is a family of stopping times $(\tau_t; t \geq 0)$ such that the map $t \to \tau_t$ is almost surely increasing, right-continuous and such that $\tau_0 = 0$ and $\lim_{t\to\infty} \tau_t = \infty$.

We will consider a *time-changed process* $(\widehat{Y}_t; t \geq 0) := (\widehat{U}_t, \widehat{V}_t; t \geq 0)$ with generator $\widehat{\mathscr{A}}$, where $\widehat{U}_t = U_{\tau_t}$ and $\widehat{V}_t = V_{\tau_t}$ for an appropriate time change $(\tau_t; t \geq 0)$. Then, by virtue of proving that the martingale problem for $(\widehat{\mathscr{A}}, \delta_0)$ is well-posed, we will show that the $(\mathscr{A}, \delta_o)$-, and so the original $(\mathscr{G}, \delta_o)$-martingale problem, has a unique solution.

Let

$$\phi_t = \int_0^t \frac{ds}{e^{2\theta V_s} - e^{-2\theta U_s}} = \int_0^t \rho(U_s, V_s) ds \ , \tag{4.22}$$

and define a family of time changes associated to ϕ_t as its left-side inverse

$$\tau_t = \inf\{s : \phi_s > t\} \ .$$

Evidently τ_t is measurable, increasing, $\tau_0 = 0$ and $\sup_t \phi_t = \infty$. The only concern is that ϕ_t can be infinite for some finite $t > 0$ which would then lead to $\lim_{t\to\infty} \tau_t < \infty$. The integrand (4.22) is almost surely finite for all $(V_s, U_s) \in (0, \infty)^2$, leaving us to deal only with the origin. First note that we can write ϕ_t in terms of the original variables as

$$\phi_t = \int_0^t \frac{ds}{e^{2\theta V_s} - e^{-2\theta U_s}} = \int_0^t \frac{e^{-\theta X_s}}{e^{\theta R_t} - e^{-\theta R_s}} ds = \int_0^t \frac{1}{2\theta} b_r(X_s, R_s) ds \ ,$$

where b_r is the R-drift coefficient of $\mathscr{G}$.

We know that there exists at least one solution to the martingale problem associated to the generator (4.7) and started at the origin, since we found an entrance law for a process started at the origin in the previous section. Recall that we can express *any* solution associated to the $(\mathscr{G}, \delta_o)$-martingale problem as a weak solution to the system of corresponding stochastic differential equations (see, for example, [69, Thm. V.20.1,

p. 160] or [42, Thm. IV.6.1, p. 201]):

$$X_t = \int_0^t d\beta_s^{(1)} ,$$
$$R_t = \int_0^t d\beta_s^{(2)} + \int_0^t b_r(X_s, R_s) ds ,$$

where $(\beta_t^{(1)}; t \geq 0)$ and $(\beta_t^{(2)}; t \geq 0)$ are two standard Brownian motions started at the origin and satisfying $\langle d\beta_t^{(1)}, d\beta_t^{(2)} \rangle = a_{xr}(X_t, R_t) = \Big(\coth(\theta R_t) - e^{-X_t\theta}\operatorname{csch}(\theta R_t)\Big)\, dt$. Thus constructed SDE's are well-defined, and we deduce that $2\theta\phi_t = \int_0^t b_r(X_s, R_s)ds < \infty$ a.s. for all $t > 0$. Thus ϕ_t is a.s. finite for all $t \geq 0$, and our proposed time change is meaningful for any solution to the $(\mathscr{G}, \delta_o)$-martingale problem. Consequently we transform our $(\mathscr{A}, \delta_o)$-martingale problem into the one associated to the generator (see [69, Ch. V.26 p. 175])

$$\widehat{\mathscr{A}^{(\theta)}} = \widehat{\mathscr{A}} = \frac{1}{2}(1 - e^{-2\theta u})\frac{\partial^2}{\partial u^2} + \theta\frac{\partial}{\partial u} + \frac{1}{2}(e^{2\theta v} - 1)\frac{\partial^2}{\partial v^2} + \theta\frac{\partial}{\partial v} \tag{4.23}$$

for $u \in [0, \infty)$, $v \in [0, \infty)$, acting on $\mathscr{C}_c([0, \infty)^2)$, the space of all bounded, infinitely continuously differentiable functions on $[0, \infty)^2$ with compact support (and vanishing at infinity and close to the axis $u = 0$ and $v = 0$). The matrix of diffusion coefficients of the above generator is degenerate as it vanishes at the origin. For this reason we cannot use theorems of Stroock and Varadhan cited before, as they do not cater for such cases. Instead we proceed in a different way: first notice that any two-dimensional process associated to the above generator can be viewed as two one-dimensional diffusions with generators

$$\widehat{\mathscr{A}_u^{(\theta)}} = \widehat{\mathscr{A}_u} = \frac{1}{2}(1 - e^{-2\theta u})\frac{\partial^2}{\partial u^2} + \theta\frac{\partial}{\partial u} , \tag{4.24}$$

and

$$\widehat{\mathscr{A}_v^{(\theta)}} = \widehat{\mathscr{A}_v} = \frac{1}{2}(e^{2\theta v} - 1)\frac{\partial^2}{\partial v^2} + \theta\frac{\partial}{\partial v} , \tag{4.25}$$

both generators acting on $\mathscr{C}_c^\infty([0, \infty))$, which is defined analogously to $\mathscr{C}_c^\infty([0, \infty)^2)$. Hence, any solution to the $(\widehat{\mathscr{A}}, \delta_{y_0})$-martingale problem, for any $y_0 = (u_0, v_0) \in [0, \infty)^2$, gives *a* solution to the $(\widehat{\mathscr{A}_u}, \delta_{u_0})$- and $(\widehat{\mathscr{A}_v}, \delta_{v_0})$-martingale problems. Our aim is to show that, conversely, every solution to $(\widehat{\mathscr{A}}, \delta_{y_0})$-martingale problem is *necessarily* a combination of solutions to the $(\widehat{\mathscr{A}_u}, \delta_{u_0})$- and $(\widehat{\mathscr{A}_v}, \delta_{v_0})$-martingale problems. This will give us an advantage of only needing to deal with one-dimensional diffusions, theory of which, unlike of higher-dimensional cases, is extensive.

Now let $(\widehat{U}, \widehat{V})$ be any solution to the $\widehat{\mathscr{A}}$-martingale problem started at the origin. Apply a stopping time $T = \inf\{t : \widehat{V}_t \geq N\}$ and consider a pair of SDE's satisfied by the process

$$\begin{aligned} \widehat{U}_t &= \int_0^{t\wedge T} \sqrt{1-e^{-2\theta \widehat{U}_s}}d\beta_s^{\hat{u}} + \int_0^{t\wedge T} \theta ds \\ \widehat{V}_t &= \int_0^{t\wedge T} \sqrt{e^{2\theta \widehat{V}_s}-1}d\beta_s^{\hat{v}} + \int_0^{t\wedge T} \theta ds \ , \end{aligned} \tag{4.26}$$

We would like to show that there is in fact a unique *strong* solution to the above system of SDE's. Thanks to the stopping, all the coefficients in the equations above are bounded. Note also that by substituting $\widehat{u}^+$, resp. $\widehat{v}^+$, instead of $\widehat{u}$, resp. $\widehat{u}$, with $x^+ := \min\{0, x\}$, we can extend the coefficients of the above SDE's continuously and boundedly from $\mathbb{R}^+$ to the whole of $\mathbb{R}$, rendering results in [42, Ch. IV] applicable in our case. Then, since coefficients of (4.26) are continuous and bounded up to the time T, there exists a weak solution to each of SDE's (4.26) started at the origin by [42, Thm. IV.2.2]. We now need to show that these solutions are pathwise, and so weakly, unique. Note that close to the origin, that is for small values of $\hat{u}$, resp. $\hat{v}$, $\widehat{U}$, resp. $\widehat{V}$, behaves like a squared Bessel process (of a certain dimension depending on the parameter θ). Knowing that the SDE for the squared Bessel process of any dimension $\alpha > 0$ has a pathwise unique solution for any starting point, we might be hopeful. Indeed, just like in the BESQ$^\alpha$ situation, a condition of Yamada and Watanabe (see, for example, [42, Thm. IV.3.2] or [69, Thm. V.40]) gives us the required uniqueness. The said Yamada-Watanabe condition requires the coefficients of an SDE to be bounded, the drift coefficient to be Lipschitz and the variance coefficient σ to satisfy a more relaxed, than the Lipschitz continuity, condition

$$|\sigma(x) - \sigma(y)| \leq \kappa(|x-y|), \quad x, y \in \mathbb{R} \tag{4.27}$$

for a strictly increasing function $\kappa : [0, \infty) \to \mathbb{R}$, such that $\kappa(0) = 0$ and $\int_{0+} \kappa^{-1}(x)dx = \infty$. In the time interval $[0, T]$ the coefficients of our SDEs are bounded and the drift coefficients, being constant, satisfy the Lipschitz condition. Now, using the inequality $|x^{1/2} - y^{1/2}| \leq |x - y|^{1/2}$, for all $x, y \geq 0$, we calculate

$$|(e^{2\theta x} - 1)^{1/2} - (e^{2\theta y} - 1)^{1/2}| \leq |e^{2\theta x} - e^{2\theta y}|^{1/2} .$$

But $f(x) := e^{2\theta x}$ is a convex function on $\mathbb{R}$ with a bounded derivative on $[0, N]$. Hence, up until the stopping (which is all we are interested in), the above is bounded above by

$$(2\theta K_N)^{1/2}|x - y|^{1/2}, \quad \text{for all } 0 \leq x, y \leq N \ ,$$

where K_N is a Lipschitz constant depending on N. Thus for $0 \leq x, y \leq N$

$$|(e^{2\theta x} - 1)^{1/2} - (e^{2\theta y} - 1)^{1/2}| \leq \kappa(|x - y|)$$

for $\kappa(x) = K|x|^{1/2}$, where K is some positive finite constant. Similarly one finds for all $|x|, |y| \leq N$

$$|(1 - e^{-2\theta x})^{1/2} - (1 - e^{-2\theta y})^{1/2}| \leq \kappa(|x - y|) \ ,$$

where κ is defined as above (but with, perhaps, a different constant K). Existence of pathwise unique solution to (4.26) issued from the origin implies the existence of the unique *strong* solution started at the origin (see, for example, [42, Thm. IV.1.1]). This is to say that there is a unique (up to indistinguishability) solution $(\widehat{U}, \widehat{V})$, which is a deterministic function of the initial condition δ_o and the Brownian motions $(\beta_t^{\hat{u}}, \beta_t^{\hat{v}}; 0 \leq t \leq T)$ (which have covariation 0). Thus, being driven by independent Brownian motions, the unique solutions to each of the SDE's (4.26) must themselves be independent processes. The uniqueness of solution to the SDE's (4.26) now implies the well-posedness of the $(\widehat{\mathscr{A}_{\hat{u}}}, \delta_o)$- and $(\widehat{\mathscr{A}_{\hat{v}}}, \delta_o)$-martingale problems ([42, Thm. IV.3.2]). But the independence of the unique solutions $\widehat{U}$ and $\widehat{V}$ means that there is essentially just one way of issuing the bivariate process $(\widehat{U}, \widehat{V})$ from the origin and consequently there is a unique solution to the $(\widehat{\mathscr{A}}, \delta_o)$-martingale problem.

Reversing the time change. What is left to do now is to translate this result back in to a result about $\mathscr{A}$, and so the original generator $\mathscr{G}$. First we show that, because the 'time-changed' $(\widehat{\mathscr{A}}, \delta_o)$-martingale problem is well-posed, the original $(\mathscr{A}, \delta_o)$-martingale problem must be well-posed as well. The well-posedness of the $(\mathscr{G}, \delta_o)$-martingale problem will then follow. We argue by contradiction. Denote by $\widehat{\mathbb{Q}}$ the unique solution to the $(\widehat{\mathscr{A}}, \delta_o)$-martingale problem and by $(\widehat{U}, \widehat{V})$ the associated process, and suppose that the $(\mathscr{A}, \delta_o)$-martingale problem has two solutions, say, $\mathbb{Q}_1$ and $\mathbb{Q}_2$. The measure $\widehat{\mathbb{Q}}$ is defined on the 'time-changed' measurable space $(\widehat{\Omega}, \{\widehat{\mathscr{F}}_t\}_{t\geq 0}, \widehat{\mathscr{F}})$ with $\mathscr{F}_t = \sigma(\widehat{U}_s, \widehat{V}_s; 0 \leq s \leq t)$, while $\mathbb{Q}_1$ and $\mathbb{Q}_2$ are defined on the 'original' measurable space $(\Omega, \{\mathscr{F}_t\}_{t\geq 0}, \mathscr{F})$ with $\mathscr{F}_t = \sigma(U_s, V_s; 0 \leq s \leq t)$. Then

$$f(U_t, V_t) - f(U_0, V_0) - \int_0^t \mathscr{A} f(U_s, V_s) ds$$

is an $\mathscr{F}_t$-martingale under $\mathbb{Q}_1$ and under $\mathbb{Q}_2$ for all $f \in \mathscr{D}_{\mathscr{A}} = \mathscr{D}_{\widehat{\mathscr{A}}} = \mathscr{C}_c^\infty([0,\infty))^2$.

Next let

$$\tau'_t = \inf\{s : \phi'_s > t\} ,$$

with $\phi'_t = \int_0^t (e^{2\theta \widehat{V}_s} - e^{-2\theta \widehat{U}_s}) ds := \int_0^t \rho'(\widehat{U}_s, \widehat{V}_s) ds$; note that $(\widehat{U}_{\tau'_s}, \widehat{V}_{\tau'_s}) = (U_s, V_s)$ for all $s \geq 0$. Now

$$f(\widehat{U}_{\tau'_t}, \widehat{V}_{\tau'_t}) - f(\widehat{U}_0, \widehat{V}_0) - \int_0^{\tau'_t} \widehat{\mathscr{A}} f(\widehat{U}_s, \widehat{V}_s) ds$$

is an $\widehat{\mathscr{F}}_{\tau'_t}$-martingale under $\widehat{\mathbb{Q}}$, the unique solution to the $(\widehat{\mathscr{A}}, \delta_o)$ martingale problem. Notice that $\mathscr{A} = \frac{1}{\rho'} \widehat{\mathscr{A}}$ and that $d\phi'_s / ds = \rho'_s$. If $f \in \mathscr{C}_c^\infty([0,\infty))^2$, then

$$\begin{aligned}\int_0^{\tau'_t} \widehat{\mathscr{A}} f(\widehat{U}_s, \widehat{V}_s) ds = \int_0^{\tau'_t} \widehat{\mathscr{A}} f(\widehat{U}_s, \widehat{V}_s) \frac{ds}{d\phi'_s} d\phi'_s = \int_0^{\tau'_t} \frac{1}{\rho'(\widehat{U}_s, \widehat{V}_s)} \widehat{\mathscr{A}} f(\widehat{U}_s, \widehat{V}_s) d\phi'_s \\ = \int_0^{\tau'_t} \mathscr{A} f(\widehat{U}_{\tau'_l}, \widehat{V}_{\tau'_l}) dl = \int_0^t \mathscr{A} f(U_l, V_l) dl \ ,\end{aligned}$$

where we have set $\phi'_s = l$ and so $\tau'_l = s$ and $(\widehat{U}_{\tau'_l}, \widehat{V}_{\tau'_l}) = (U_l, V_l)$, $l \geq 0$. Hence,

$$f(\widehat{U}_{\tau'_t}, \widehat{V}_{\tau'_t}) - f(\widehat{U}_0, \widehat{V}_0) - \int_0^{\tau'_t} \widehat{\mathscr{A}} f(\widehat{U}_s, \widehat{V}_s) ds$$
$$= f(U_t, V_t) - f(U_0, V_0) - \int_0^t \mathscr{A} f(U_s, V_s) ds \; .$$

The RHS, and so the LHS, must be a martingale under $\mathbb{Q}_1$ and $\mathbb{Q}_2$ for all $f \in \mathscr{D}_{\mathscr{A}}$ with respect to $\sigma(U_s, V_s; 0 \leq s \leq t) = \sigma(\widehat{U}_{\tau'_s}, \widehat{V}_{\tau'_s}; 0 \leq \tau'_s \leq \tau'_t)$. But $\widehat{\mathbb{Q}}$ is the only measure under which the LHS is a martingale with respect to the natural filtration of $(\widehat{U}, \widehat{V})$ for all $f \in \mathscr{D}_{\mathscr{A}}$; we thus see that, intuitively, measures $\mathbb{Q}_1$ and $\mathbb{Q}_2$ on $(\Omega, \{\mathscr{F}_t\}_{t\geq 0}, \mathscr{F})$ after the time change correspond to the same measure $\widehat{\mathbb{Q}}$ on $(\widehat{\Omega}, \{\widehat{\mathscr{F}}_t\}_{t\geq 0}, \widehat{\mathscr{F}})$. This means that $\mathbb{Q}_1 = \mathbb{Q}_2$ and that, consequently, the $(\mathscr{A}, \delta_o)$-martingale problem is well-posed. Now, since the $(\mathscr{A}, \delta_o)$-martingale problem is well-posed, so is our original problem associated to $(\mathscr{G}, \delta_o)$. We denote the unique solution to the $(\mathscr{G}, \delta_o)$-martingale problem by $\mathbb{P}_o$.

□

Drifting case, $\mu > 0$. Finally we prove

Proposition 4.18. *The martingale problem for $\mathscr{G}^\mu$, for $\mu > 0$, started at the origin is well-posed.*

Proof. The beginning of the proof essentially mimics the proof of Proposition 4.17 First of all we apply the coordinate transformation (4.9), thus obtaining the generator in terms of the new variables $(u, v) = y$

$$\mathscr{A}^{(\theta,\mu)} = \mathscr{A}^\mu = \frac{1 - e^{-2\theta u}}{2(e^{2\theta v} - e^{-2\theta u})} \frac{\partial^2}{\partial u^2} + \frac{\theta + \mu(e^{-2\theta u} - 1)}{e^{2\theta v} - e^{-2\theta u}} \frac{\partial}{\partial u}$$
$$+ \frac{e^{2\theta v} - 1}{2(e^{2\theta v} - e^{-2\theta u})} \frac{\partial^2}{\partial v^2} + \frac{\theta + \mu(e^{2\theta v} - 1)}{e^{2\theta v} - e^{-2\theta u}} \frac{\partial}{\partial v}$$

with the domain $\mathscr{C}_c^\infty([0, \infty)^2)$.

To solve the $(\mathscr{G}^\mu, \delta_o)$-martingale problem we will use the change of measure argument just like we did in the case of starting away from the boundary. However, explosion of the R-drift at o makes the time change once again necessary. In fact we use the same time change as before: $(\tau_t; t \geq 0)$ with $\tau_t = \inf\{s : \phi_s > t\}$ and

$$\phi_t = \int_0^t \rho(U_s, V_s) ds = \int_0^t \frac{ds}{e^{4\theta V_s} - e^{-4\theta U_s}} = \frac{1}{2\theta} \int_0^t \left(b_r^\mu(X_s, R_s) - \mu a_{xr}^\mu(X_s, R_s) \right) ds \ ,$$

where b_r^μ and a_{xr}^μ are coefficients of $\mathscr{G}^\mu$ (4.1). We proceed to show that $\phi_t < \infty$ for all $t \geq 0$. Once again notice that to each solution of the $\mathscr{G}^\mu$-martingale problem started at the origin (again, from Corollary 4.16 we know there exists at least one) there corresponds a weak solution to the following system of SDE's

$$\begin{aligned} X_t &= \int_0^t d\beta_s^{(1)} + \mu t \ , \\ R_t &= \int_0^t d\beta_s^{(2)} + \int_0^t b_r^\mu(X_s, R_s) ds \ , \end{aligned} \tag{4.28}$$

where $b_r^\mu(x,r) = \mu(\coth(\theta r) - e^{-x\theta}\operatorname{csch}(\theta r)) + \theta e^{-\theta x}\operatorname{csch}(\theta r)$, for $(x,r) \in W$, and $(\beta_t^{(1)}; t \geq 0)$ and $(\beta_t^{(2)}; t \geq 0)$ are two standard Brownian motions started at zero and satisfying $\langle d\beta_t^{(1)}, d\beta_t^{(2)} \rangle = a_{xr}^\mu(X_t, R_t) = (\coth(\theta R_s) - e^{-X_s\theta}\operatorname{csch}(\theta R_s))dt$ for all $t \geq 0$. Since the SDE's above are well-defined, we have $\int_0^t b_r^\mu(X_s, R_s) ds < \infty$ a.s. for all $t > 0$. Consequently, using the fact that $|a_{xr}^\mu(x,r)| < 1$ for all $(x,r) \in W'$, we deduce

$$\phi_t = \frac{1}{2\theta} \int_0^t \left(b_r^\mu(X_s, R_s) - \mu a_{xr}^\mu(X_s, R_s) \right) ds < \infty \quad \text{a.s. for all } t > 0 \ .$$

Thus, since $\phi_t < \infty$ for all $t > 0$, we have a well-defined time-change which is meaningful for any solution to the $(\mathscr{A}^{(\theta,\mu)}, \delta_o)$-martingale problem; the generator associated to the time-changed process is given by

$$\begin{aligned} \widehat{\mathscr{A}_\mu^{(\theta)}} = \widehat{\mathscr{A}_\mu} = \frac{1}{2}(1 - e^{-2\theta u}) \frac{\partial^2}{\partial u^2} + (\theta + \mu(e^{-2\theta u} - 1)) \frac{\partial}{\partial u} \\ + \frac{1}{2}(e^{2\theta v} - 1) \frac{\partial^2}{\partial v^2} + (\theta + \mu(e^{2\theta v} - 1)) \frac{\partial}{\partial v} \ . \end{aligned}$$

Now, notice that, if we denote by $\hat{a}$ and $\hat{b}$ the coefficients of the generator $\widehat{\mathscr{A}}$ (4.23), then the coefficients of $\widehat{\mathscr{A}^\mu}$ are given by $\hat{a}$ and $\hat{a} + c\hat{b}$ with $c = 2\mu(-1,1)$. If we consider the system up to the stopping time $T = \inf\{t : \widehat{V} \geq N\}$ for some $N \in \mathbb{R}$, then all the coefficients of the generator $\widehat{\mathscr{A}}$ are bounded and so is the process $\langle c, ac \rangle$. Then, since the $(\widehat{\mathscr{A}}, \delta_o)$-martingale problem is uniquely posed, by Theorem 4.8 there is a unique solution to the $\widehat{\mathscr{A}^\mu}$-problem started at the origin, at least up to the stopping time T. However, we know that, started away from the origin, the $\widehat{\mathscr{A}^\mu}$-martingale problem is well-posed (since by Proposition 4.7 so is the $(\mathscr{G}^\mu, \delta_{z_0})$-martingale problem for all $z_0 \in W'$). Hence, there is only one possible process associated to the generator $\widehat{\mathscr{A}^\mu}$ and issued from the origin. By the same 'reversing the time-change' argument, similar to the one at the end of the proof of Proposition 4.17, this implies that there is a unique solution to the $(\mathscr{A}^\mu, \delta_o)$- and $(\mathscr{G}^\mu, \delta_o)$-martingale problems.

□

We conclude that the $\mathscr{G}_\mu$, for $\mu \geq 0$, started at the origin has a unique solution and, as a consequence, the entrance law from Corollary 4.16 for the process (X, R) started at the origin is unique.

4.5 Concluding remarks

Suppose $\mathscr{L}$ is an infinitesimal generator acting on $\mathscr{C}_c^\infty(W)$ and such that the $(\mathscr{L}, \Lambda^\infty(r_0, \cdot))$-martingale problem is well-posed for all $r_0 > 0$ with the solution hitting the boundary ∂W with probability 0. Let $(P_t; t > 0)$ be the semi-group of the associated process. From our analysis in section 4.2 it follows that for $(Q^\mu \Lambda^\mu)$ to solve the $(\mathscr{L}, \Lambda^\mu(r_0, \cdot))$-forward equation it is *sufficient* for the coefficients of the generator $\mathscr{L}$ to satisfy

$$b_x = \mu \ ,$$

$$a_{xx} = a_{rr} = 1 \ ,$$

$$\partial_x a_{xr} = b_r - \mu a_{xr} \ .$$

It follows that the above conditions on the generator's coefficients are also sufficient for the intertwining $P_t \Lambda^\mu = \Lambda^\mu Q_t^\mu$ to hold and, consequently, for the marginal processes to be distributed as BM(μ) and BES^3(μ).

There are several interesting question which remain to be answered: are there examples of two-dimensional diffusions in the wedge W with the $(\mathrm{BM}(\mu), \mathrm{dBES}^3(\mu))$ marginals, $\mu \geq 0$, which do not satisfy neither the intertwining nor the Dynkin criteria? We have seen an example of such process in the discrete setting so a likely answer is yes. An even harder question still is: can we describe *all* the bi-variate diffusions in the wedge W with the $(\mathrm{BM}, \mathrm{BES}^3(\mu))$ marginals?

Chapter 5

Process in the Gelfand-Cetlin cone

The main subject of this chapter is the so-called Hermitian minor process, a joint process of eigenvalues of consecutive principal minors of the Hermitian Brownian motion, which can be seen as a generalisation of the $\theta = 0$ coupling $Z^{(\theta)} = (\mathrm{BM}, \mathrm{BES}^3)$ discussed in Chapter 4. In Chapter 3 we saw an example of the simplest (non-trivial) GUE minor process consisting of the eigenvalues of the first 2 principal minors of a 2×2 Hermitian Brownian motion. By looking at the eigenvalues of the principal minors of GUE matrices of higher dimensions, one can identify a more complex process, that we call Hermitian minor process, taking values in the continuous Gelfand-Cetlin cone, but which, unlike in the 2-dimensional case, is not Markov. However, we prove that the joint process of the eigenvalues of the first two principal minors *is* in fact Markov and find its generator. Moreover, we make a first step towards identifying a discrete analogue of this process as a joint process of the top two rows of a certain Markov chain $(\mathscr{X}(k); k \geq 0)_q$ living in the discrete Gelfan-Cetlin cone, when $q = 1$. The family of Markov processes $(\mathscr{X}(k); k \geq 0)_q$ constitute an higher-dimensional generalisation of the family of Markov chains arising in representation theory of $U({}_q\mathfrak{sl}_2)$ described in Chapter 3. We present an informal argument showing that the generator of appropriately scaled top two rows of $\mathscr{X}$, for $q = 1$, converge to the generator of the process generated by the top two rows of the GUE minor process.

Also in this chapter we review the $q = 0$ ($\theta = \infty$, in the continuous setting) part of the story in higher dimensions. In particular in Section 5.2 we describe a certain transformation for random walks introduced by O'Connell and Yor [60], which leads to a higher dimensional generalisation of the Pitman's theorem and explain how these results are related to our work in Chapter 3. The unifying factor for process described in Section 5.2 and the Hermitian minor process is the fact that the top layer of both processes evolves as an n-point Dyson's Brownian motion. However, while the process of O'Connell and Yor has a discrete version, in fact it is *defined* as a diffusion limit of the discrete chain constructed via the Robinson-Schensted algorithm, the Hermitian minor process, for $n \geq 3$, doesn't have a discrete counterpart. One of the main achievements of this chapter is making a step towards identifying it; this is done in section 5.4.

5.1 Dyson's Brownian motions and the eigenvalues of random matrices

In this section we will review the notion of *non-colliding Brownian motions*. We will show how these can be constructed as an h-transform of an n-dimensional Brownian motion killed when any of its components collide. Then we will see that this same process arises as the eigenvalue process of an $n \times n$ Hermitian Brownian motion.

5.1.1 H-transformed Brownian motions and non-colliding Brownian motions.

Let B be the standard n-dimensional Brownian motion started away from the origin, defined on some filtered probability space $(\Omega, \{\mathscr{F}_t\}_{t\geq 0}, \mathscr{F}, \mathbb{P})$, and let $h(x) = \triangle(x)$, where

$$\triangle(x) = \prod_{i<j}(x_i - x_j), \quad x \in \mathbb{R}^n ,$$

is the Vandermonde determinant. We call the *Weyl chamber* a subset of $\mathbb{R}^n$

$$\mathbb{W} = \mathbb{W}_n = \{x \in \mathbb{R}^n : x_1 > x_2 > ... > x_n\} .$$

One can check that the Vandermonde determinant is harmonic with respect to the generator of an n-dimensional Brownian motion (see [53, Lemma 3.1]). Let $B_0 = x \in \mathbb{W}$ and denote by T the exit time of B from the Weyl chamber, that is $T = \inf\{t : B_t \notin \mathbb{W}\}$. Then we can define a new measure

$$\mathbb{Q}_x(A) = \mathbb{E}_x\left[\frac{h(X_t)}{h(x)}\mathbf{1}_{\{A,t<T\}}\right], \quad A \in \mathscr{F}_t\ ,$$

where $\mathbb{P}_x$ is the law of B started at x. The process in $\mathbb{W}$ with measure $\mathbb{Q}$ is called the h-transform of the Brownian motion killed when it leaves the Weyl chamber. Observing that $\frac{\partial}{\partial x_i}log(\triangle(x)) = \frac{1}{\triangle(x)}\frac{\partial}{\partial x_i}\triangle(x)$ and using (3.9) one calculates the generator of the transformed process to be

$$\mathscr{G}_h = \frac{1}{2}\sum_i \frac{\partial^2}{\partial x_i^2} + \sum_i \Big[\sum_{j\neq i}\frac{1}{x_i - x_j}\Big]\frac{\partial}{\partial x_i}\ . \tag{5.1}$$

It is possible to start the process at the origin: if $\mathbb{Q}_x$ denotes the law of the transformed process started at $x \in \mathbb{W}$, then the limit $\lim_{\substack{x\to 0\\ x\in\mathbb{W}}} \mathbb{Q}_x$ is well-defined (see [53, Lemma 3.2]). Indeed, one notices that for $n = 2$ the generator (5.1) corresponds to the SDE's describing the two eigenvalues of the 2×2 Hermitian Brownian motion (see Section 3.1.1 and Section 5.1.2 below), which, we know, are issued from the origin. We will generalise this in Section 5.1.2.

Brownian motion in the Weyl chamber can be interpreted as *Brownian motions conditioned not to collide*. Of course, the event of n independent Brownian motions started at some point $x \in \mathbb{W}$ not ever colliding has probability 0, so this statement needs some explanation. If T is the first time B leaves the Weyl chamber $\mathbb{W}$, then the probability $\mathbb{P}_x(B_s \in dx|t < T)$ for all $t, s > 0$ is well defined. One can then let t tend to infinity. The limiting process will be Brownian motion conditioned not to exit the Weyl chamber.

5.1.2 Dyson's Brownian motions on the spectra of Hermitian Brownian motion

In this section we describe a natural context in which non-colliding Brownian motions arise; we come back to the discussion of the eigenvalues of Hermitian Brownian motion that we have started in Chapter 3. In Section 3.1.1 we studied the 2×2 Hermitian Brownian motion and derived the SDE's satisfied by its two eigenvalues. We now consider the general $n\times n$ case. Freeman Dyson [28] was the first to identify eigenvalues of Hermitian Brownian motion as a system of n repelling Brownian motions and derive a system of SDE's that they satisfy. It is in his honour that Brownian motions conditioned not to collide are often called *Dyson's Brownian motions*.

Theorem 5.1. *(Dyson [28]) The eigenvalues $\lambda=(\lambda_1(t),...,\lambda_n(t);t\geq 0)$ of an $n\times n$ Hermitian Brownian motion started at the origin satisfy the following stochastic differential equations*

$$d\lambda_k(t)=d\beta_k(t)+\sum_{l\neq k}\int_0^t \frac{1}{\lambda_k(s)-\lambda_l(s)}ds, \quad 1\leq k\leq n\ , \tag{5.2}$$

where $\beta=(\beta_1(t),...,\beta_n(t),t\geq 0)$ is a standard n-dimensional Brownian motion started at the origin. In particular, $d\langle\beta_k,\beta_l\rangle=\delta_{kl}dt$.

We include a proof of Dyson's result because it involves derivation of an expression for the Brownian motions driving eigenvalues $\lambda=(\lambda_1,...,\lambda_n)$ in terms of the entries of the matrix and the eigenvectors corresponding to λ in our chosen coordinate system. This will then be used in the proof of Theorem 5.2.

Proof. The proof is in the spirit of proof of [49, Thm. 10]. Let $H=(H(t),t\geq 0)$ be a Hermitian Brownian motion with $H_{ij}=(h_{ij}+ih_{ji})/\sqrt{2}$, for $i<j$, and $H_{ii}=h_{ii}$, where h_{ij}'s are independent standard Brownian motions. By $sp(H(t))$ we denote the spectrum of $H(t)$. The eigenvalue process $\lambda=(\lambda_1(t),...,\lambda_n(t);t\geq 0)$ is a function of the n^2 independent Brownian motions constituting the entries of H. In order to apply Itô's lemma to λ, we need to calculate first- and second-order partial derivatives of λ with respect to the coordinates of the Brownian motions driving H. For any $t>0$, any $\lambda(t)\in sp(H(t))$ satisfies $\det|H(t)-\mathbb{I}\lambda(t)|=0$, where $\mathbb{I}$ is the $n\times n$ identity matrix.

However, calculating the required quantities directly from this identity is not feasible. Instead we use the fact that Hermitian matrices are invariant under the conjugation by the Unitary matrices. In what follows we suppress dependence of H and λ on t to simplify the notation.

Let $\mathscr{U}_n$ and $\mathscr{H}_n$ be the spaces of all $n \times n$ Unitary and Hermitian matrices respectively. Then the map $\varphi_U : \mathscr{H}_n \to \mathscr{H}_n$ defined by

$$\varphi_U(H) = U^* H U, \qquad U \in \mathscr{U}_n \ ,$$

is an automorphism. Moreover, the spectrum of $\varphi(H)$ is the same as the spectrum of H. If U, for example, is taken to be a matrix of normalised eigenvectors of H (i.e. $u_i = (u_{1i}, ..., u_{ni})$, the i^{th} column of U, is an eigenvector corresponding to λ_i) and Λ is a diagonal matrix of the corresponding eigenvalues, then $\varphi_U = U^* H U = \Lambda$. We say that two Hermitian matrices H and $\widetilde{H}$ are *equivalent* if $\widetilde{H} = U^* H U$ for some $U \in \mathscr{U}_n$.

If for any $\lambda \in sp(H)$ we write λ as a function of H, i.e. as a map $\lambda : \mathbb{R}^{n^2} \to \mathbb{W}_n$, then $\lambda(H) = \lambda(\widetilde{H})$ for any $\widetilde{H}$ equivalent to H. A simple application of a chain rule then gives

$$\frac{\partial \lambda(H)}{\partial h_{ij}} = \frac{\partial \lambda(\varphi_U(H))}{\partial h_{ij}} = \sum_{r,s} \frac{\partial \lambda(\widetilde{H})}{\partial \tilde{h}_{rs}} \frac{\tilde{h}_{rs}}{\partial h_{ij}} \tag{5.3}$$

and

$$\frac{\partial^2 \lambda(H)}{\partial h_{ij} \partial h_{lp}} = \sum_{r,s,k,m} \frac{\partial^2 \lambda(\widetilde{H})}{\partial \tilde{h}_{rs} \tilde{h}_{km}} \frac{\partial \tilde{h}_{rs}}{\partial h_{ij}} \frac{\partial \tilde{h}_{km}}{\partial h_{lp}} + \sum_{l,p} \frac{\lambda(\widetilde{H})}{\partial \tilde{h}_{rs}} \frac{\partial^2 \tilde{h}_{rs}}{\partial h_{ij}^2} \tag{5.4}$$

where $\tilde{h}_{ij}$'s denote Brownian motions constituting entries of $\widetilde{H}$. The above identities hold true for any $U \in \mathscr{U}_n$ such that $\widetilde{H} = U^* H U$, and, in particular, for U chosen such that $\widetilde{H} = \Lambda = \operatorname{diag}(\lambda)$.

Now, for any $\widetilde{H}$ equivalent to H and $\lambda \in \operatorname{sp}(H)$ we have

$$0 = \det|\widetilde{H} - \lambda \mathbb{I}| = \sum_{\sigma \in \mathbf{S}_n} \prod_{i=1}^{n} (\widetilde{H}_{i\sigma(i)} - \delta_{i\sigma(i)} \lambda) \ , \tag{5.5}$$

where $\mathbf{S}_n$ is the n^{th} symmetric group. Implicit differentiation of the RHS of (5.5) at

$\widetilde{H} = \Lambda$ and rearranging gives

$$\left.\frac{\partial \lambda_j}{\partial \tilde{h}_{ii}}\right|_{\widetilde{H}=\Lambda} = \delta_{ij} \quad \text{and} \quad \left.\frac{\partial \lambda_k}{\tilde{h}_{ij}}\right|_{\widetilde{H}=\Lambda} = 0, \ \forall i \neq j \neq k \ . \tag{5.6}$$

Next, differentiating (5.5) twice with respect to $\tilde{h}_{ij}$, $i < j$, and letting $\widetilde{H} = \Lambda$ yields

$$\left.\frac{\partial^2 \lambda_i}{\partial \tilde{h}_{ij}^2}\right|_{\widetilde{H}=\Lambda} = \left.\frac{\partial^2 \lambda_i}{\partial \tilde{h}_{ji}^2}\right|_{\widetilde{H}=\Lambda} = -\frac{\prod_{k \neq i,j}(\lambda_k - \lambda_i)}{\sum_p \prod_{k \neq p}(\lambda_k - \lambda_i)} = \frac{1}{\lambda_i - \lambda_j} \tag{5.7}$$

and $\frac{\partial^2 \lambda_j}{\partial \tilde{h}_{ij}^2} = -\frac{\partial^2 \lambda_j}{\partial \tilde{h}_{ji}^2}$. All the other second order partial derivatives are zero.

To find derivatives of the form $\partial \tilde{h}_{lp}/\partial h_{ij}$, we use the expression

$$\widetilde{H}_{pl} = \sum_{i,j} u_{ip}^* H_{ij} u_{jl}$$

and note that $\widetilde{H}_{ii} = \tilde{h}_{ii}$, $\tilde{h}_{lp} = \frac{1}{\sqrt{2}}(\widetilde{H}_{pl} + \widetilde{H}_{lp})$ and $\tilde{h}_{pl} = \frac{1}{\sqrt{2}}(\widetilde{H}_{pl} - \widetilde{H}_{lp})i$, for $i < j$. So, for example, for $l < p$ and $i < j$

$$\frac{\partial \tilde{h}_{lp}}{\partial h_{ij}} = \frac{1}{\sqrt{2}}(u_{il}^* u_{jp} + u_{jl}^* u_{ip} + u_{ip}^* u_{jl} + u_{jp}^* u_{il}) \ .$$

Evidently, since the transformation φ taking H to $\widetilde{H}$ is linear, all the second-order derivatives are zero. Combining this with (5.6) and (5.7) we can now evaluate the right-hand sides of equations (5.3) and (5.4). We calculate, for $1 \leq k \leq n$

$$\frac{\lambda_k}{\partial h_{ij}} = \frac{1}{\sqrt{2}}(u_{ik}^* u_{jk} + u_{jk}^* u_{ik}), \quad \frac{\lambda_k}{\partial h_{ji}} = \frac{1}{\sqrt{2}}(u_{ik}^* u_{jk} - u_{jk}^* u_{ik})i, \quad i < j \ ,$$
$$\frac{\lambda_k}{\partial h_{ii}} = u_{ik}^* u_{ik}, \quad 1 \leq i \leq n \ .$$

Thus, finally applying Itô's lemma, we find the process driving the k^{th} eigenvalue to satisfy

$$d\beta_k = \sum_{i,j} \frac{\partial \lambda_k}{\partial h_{ij}} dh_{ij} = \sum_{i<j} \left[\frac{1}{\sqrt{2}} (u_{ik}^* u_{jk} + u_{jk}^* u_{ik}) dh_{ij} + \right.$$
$$\left. + \frac{1}{\sqrt{2}} (u_{ik}^* u_{jk} - u_{jk}^* u_{ik}) i dh_{ji} \right] + \sum_i u_{ik}^* u_{ik} dh_{ii}$$
$$= \sum_{i,j} u_{ik}^* u_{jk} dH_{ij} \ , \quad (5.8)$$

where $u_k = (u_{1k}, ..., u_{nk})$ is the normalised eigenvector associated to the eigenvalue λ_k.

Using orthonormality of eigenvectors u, i.e. the fact that $\sum_k u_{ki}^* u_{kj} = \delta_{ij}$, one easily shows that

$$d\langle \beta_k(t), \beta_m(t) \rangle = \delta_{km} dt \ .$$

It follows by Lévy's characterisation theorem that $(\beta_1, ..., \beta_n)$ is a standard n-dimensional Brownian motion, and that, in particular, each eigenvalue is driven by a Brownian motion independent of the driving Brownian motions of all the other eigenvalues.

Now substituting (5.6) and (5.7) into (5.4) we have

$$\sum_{r,s,i,j} \frac{\partial^2 \lambda_k}{\partial h_{ij} \partial h_{r,s}} = \sum_{i,j} \frac{\partial^2 \lambda_k}{\partial h_{ij}^2} = \sum_{i,j} \sum_{l \neq k} \left[\frac{\partial^2 \lambda_k}{\partial \tilde{h}_{lk}^2} \times \left(\frac{\partial \tilde{h}_{lk}}{\partial h_{ij}} \right)^2 + \frac{\partial^2 \lambda_k}{\partial \tilde{h}_{kl}^2} \times \left(\frac{\partial \tilde{h}_{kl}}{\partial h_{ij}} \right)^2 \right] \Bigg|_{\tilde{H}=\Lambda}$$
$$= \sum_{l \neq k} \frac{1}{\lambda_k - \lambda_l} \sum_{i,j} \left[\left(\frac{\partial \tilde{h}_{lk}}{\partial h_{ij}} \right)^2 + \left(\frac{\partial \tilde{h}_{kl}}{\partial h_{ij}} \right)^2 \right] \Bigg|_{\tilde{H}=\Lambda}$$
$$= \sum_{l \neq k} \frac{1}{\lambda_k - \lambda_l} \ ,$$

where the last equality follows by using explicit expressions for derivatives $\partial \tilde{h}_{lk} / h_{ij}$ and the unitarity of U. This completes the proof.

$\square$

5.1.3 A Markov function of the GUE minor process

In Section 3.1.1 we considered a joint process of the eigenvalues of a 2×2 Hermitian Brownian motion together with the eigenvalue of its first principal minor. In this section we generalise this construction. Let H be a GUE matrix. For all $0 \leq m \leq n-1$ denote by H_{n-m} the m^{th} principal minor of H, i.e. $H_{n-m} = \{H_{ij}\}_{1 \leq i,j \leq n-m}$. Note that $H_n = H$. If $\lambda = (\lambda_1, ..., \lambda_n)$ and $\mu = (\mu_1, ..., \mu_{n-1})$ are ordered eigenvalues of H and its first principal minor correspondingly, then by the Cauchy's interlacing theorem, also called Rayleigh theorem, $\lambda_n \leq \mu_{n-1} \leq ... \leq \mu_1 \leq \lambda_1$, i.e. $\mu \preceq \lambda$. It follows by induction that the eigenvalues of all the principal minors of H form a pattern in the (continuous) Gelfand-Cetlin cone of depth n

$$\mathbb{K}_n = \{(x^n, x^{n-1} ..., x^1) \in \mathbb{R}^n \times \mathbb{R}^{n-1} \times ... \times \mathbb{R} : x^n \succeq x^{n-1} \succeq ... \succeq x^1\} \ .$$

In [47] Johansson and Nordenstam prove that, not only that the joint measure of the eigenvalues of a GUE matrix is determinantal, but that also so is the joint measure of all the eigenvalues of the GUE principal minors and compute the associated correlation functions explicitly. The authors also identify the joint distribution of the eigenvalues of principal minors of a GUE matrix as a limit of several tiling models, a polynuclear growth model and a model with RSK algorithm dynamics.

In [6] Baryshnikov proves that, given the eigenvalues of the main matrix H, i.e. the top row of the corresponding minor process, the eigenvalues of all the principal minors are distributed uniformly in $\mathbb{K}_\lambda := \{x \in \mathbb{K}_n : x^n = \lambda\}$. Let $B = (B_1, ..., B_n)$ be a standard n-dimensional Brownian motion. Baryshnikov also proves, by considering GC patterns associated to a random Young tableau as its size tends to infinity, that the largest eigenvalue of H_{n-m}, $0 \leq m \leq n-1$, is distributed as a variable

$$\Gamma_{n-m}^{(n-m)}(B(1)) := \sup_{0 \leq t_1 \leq ... \leq t_{n-m} = 1} \sum_{k=1}^{n-m} \{B_k(t_k) - B_k(t_{k-1})\} \ .$$

In particular, the largest eigenvalue of a GUE matrix H has the same law as $\Gamma_n^{(n)}(B(1))$. (The transformations $\Gamma^{(m)}$ will be defined in Section 5.2 of the chapter.)

Now, let $H = (H(t), t \geq 0)$ be an $n \times n$ Hermitian Brownian motion. Then the eigenvalues of all the principal minors of H form a continuous-time process in $\mathbb{K}_n$. We call this process the *minor process of Hermitian Brownian motion* or simply *Hermitian minor process*. It is the minor process that is the focus of our attention in this section.

In Chapter 3 we have seen that the joint process of the eigenvalues of the principal minors of a 2×2 Hermitian Brownian motion is a diffusion; this is also true in the trivial 1×1 case. In general, for $n \geq 3$, however, this is not true [1]. Deffouseux [25, Remark 11.1] remarks that this fact must have a quantum probability interpretation. On the other hand, each individual row of the minor process forms a Markov process by construction. In particular, the m^{th} row evolves as an $(n-m+1)$-point Dyson's Brownian motion. In this section we show that the process formed by the *top two* rows is a diffusion also. The main result is

Theorem 5.2. *Let $\lambda = (\lambda_1(t), ..., \lambda_n(t); t \geq 0)$ and $\mu = (\mu_1(t), ..., \mu_{n-1}(t); t \geq 0)$ be the first and the second rows of the GUE minor process respectively. Then $(\lambda(t), \mu(t); t \geq 0)$ is a diffusion with values in $\mathbb{W}_{n,n-1} := \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^{n-1} : y \preceq x\}$ and the infinitesimal generator*

$$\mathscr{G} = \frac{1}{2}\sum_i \frac{\partial^2}{\partial \lambda_i^2} + \frac{1}{2}\sum_j \frac{\partial^2}{\partial \mu_j^2} + \sum_i \sum_{k \neq i} \frac{1}{\lambda_i - \lambda_k}\frac{\partial}{\partial \lambda_i} + \sum_j \sum_{k \neq j} \frac{1}{\mu_j - \mu_k}\frac{\partial}{\partial \mu_j} + \sum_{i,j} \frac{\gamma_j^2}{(\lambda_i - \mu_j)} \frac{\prod_{\substack{1 \leq k \leq n-1 \\ k \neq j}} (\lambda_i - \mu_k)}{\prod_{\substack{1 \leq k \leq n \\ k \neq i}} (\lambda_i - \lambda_k)} \frac{\partial^2}{\partial \lambda_i \partial \mu_j}, \quad (5.9)$$

where

$$\gamma_j^2 = \left(\sum_{i=1}^n \frac{1}{(\lambda_i - \mu_j)} \frac{\prod_{\substack{1 \leq k \leq n-1 \\ k \neq j}} (\lambda_i - \mu_k)}{\prod_{\substack{1 \leq k \leq n \\ k \neq i}} (\lambda_i - \lambda_k)} \right)^{-1}. \quad (5.10)$$

That is, we claim that the eigenvalues of the first two consecutive principal minors of Hermitian Brownian motion form a Markov process of two interlacing Dyson Brownian motions with a specific interaction term between the Brownian motions from different layers. A different example of a Markov process of two interlaced Dyson Brownian motions can be found in a paper by Warren [79].

Remark. Results stated in Theorem 5.2 were also proved independently by Adler, Nordenstam and van Moebeke [1].

We also state

Conjecture 5.3. *The martingale problem associated to the generator $\mathscr{G}$ and started at $v_0 \in \mathbb{W}'_{n,n-1}$, the interior of $\mathbb{W}_{n,n-1}$, or the origin o is well-posed.*

Proof of Theorem 5.2. The diffusion and the drift terms of the generator follow directly from the fact that both the top and the second rows of the minor process evolve as n- and $(n-1)$-point Dyson's Brownian motions respectively. The difficulty lies in calculating the mixed terms of the generator. To achieve this goal we follow closely the exposition of Baryshnikov [6, Sec. 4], who notes that rotational invariance of Hermitian Brownian motion means that instead of the first principal minor of H we can consider a restriction of H to a certain random $(n-1)$-dimensional subspace of $\mathbb{C}^n$.

Let $H = (H(t); t \geq 0)$ be a Hermitian Brownian motion with the eigenvalue process $\lambda = (\lambda_1(t), ..., \lambda_n(t); t \geq 0)$ and the corresponding eigenvectors $(u_1(t), ..., u_n(t); t \geq 0)$. For any $t > 0$, $H(t)$ defines a Hermitian form on $\mathbb{C}^n$. From now on we fix $t > 0$ and suppress an explicit dependence on time to simplify the notation. Let $L \simeq \mathbb{C}^{n-1}$ be a random hyperplane defined by $L = \{x \in \mathbb{C}^n : \langle x, l \rangle = 0\}$, where $l \in \mathbb{C}^n$ is a vector distributed uniformly in the unit sphere $\mathbf{S}_n = \{x \in \mathbb{C}^n : \langle x, x \rangle = 1\}$. Let $\widetilde{H}$ be a restriction of H to L, i.e.

$$\widetilde{H}x = P_L H x, \quad x \in L \,, \tag{5.11}$$

where $P_L x = x - \langle x, l \rangle l$, $x \in \mathbb{C}^n$, is an orthogonal projection on L. In particular $x^* \widetilde{H} x = x^* H x$ for all $x \in L$. The restriction $\widetilde{H}$ to an $(n-1)$-dimensional subspace L has n eigenvalues, one and only one of which is 0; let $\mu = (\mu_1, ..., \mu_{n-1})$ denote $\widetilde{H}$'s $n-1$ non-zero eigenvalues and $\{v_1, ..., v_{n-1}\}$ – the corresponding normalised eigenvectors. Note that $\{v_1, ..., v_{n-1}\}$ forms an orthonormal basis of L and that we can write $P_L = \sum_i v_i v_i^*$. Then

$$\widetilde{H} = P_L H P_L \,,$$

and so

$$\widetilde{H}_{ij} = \sum_{k,l} p_{ik} H_{kl} p_{lj} \,, \tag{5.12}$$

where p_{ik} is the ik^{th} entry of P_L. One sees that thus defined $\widetilde{H}$ is again a Hermitian operator and agrees with (5.11). Note that if we take l to be the n^{th} unit vector, then $\widetilde{H}$ is just H's first principal minor.

We will need the following lemma

Lemma 5.4. *Let $\lambda_i \in sp(H)$ and $\mu_j \in sp(\widetilde{H})$ and let α_i and β_j be driving Brownian motions of λ_i and μ_j respectively. Then*

$$d\langle \alpha_i(t), \beta_j(t)\rangle = |\langle u_i(t), v_j(t)\rangle|^2 dt \,, \tag{5.13}$$

where u_i, resp. v_j, is the normalised eigenvector of λ_i, resp. μ_j.

Proof. Equation (5.8) in the proof of Theorem 5.1 gives a representation of Brownian motion α_i, resp. β_j, in terms of the eigenvector u_i, resp. v_j, and the entries of the matrix H, resp. $\widetilde{H}$

$$d\alpha_i = \sum_{l,r} u_{li}^* u_{ri} dH_{lr} \,,$$
$$d\beta_j = \sum_{k,m} v_{kj}^* v_{mj} d\widetilde{H}_{km} \,,$$

where, as before, we have $u_i = (u_{1i}, ..., u_{ni})$ and, similarly, $v_j = (v_{1j}, ..., v_{nj})$, for all $1 \le i, j \le n$. Hence,

$$d\langle \alpha_i, \beta_j\rangle = \sum_{l,r} u_{li}^* u_{ri} \sum_{k,m} v_{kj}^* v_{mj} d\langle H_{lr}, \widetilde{H}_{km}\rangle \tag{5.14}$$

At the same time equation (5.12) and the fact that $d\langle H_{ij}, H_{lr}\rangle = \delta_{il}\delta_{jr} dt$ give $d\langle H_{lr}, \widetilde{H}_{km}\rangle = \sum_s v_{ks} v_{rs}^* \sum_t v_{lt} v_{mt}^* dt$. Using this and orthonormality of v_i's, (5.14) becomes

$$d\langle \alpha_i(t), \beta_j(t)\rangle = \sum_{l,r} u_{li}^* u_{ri} \sum_{k,m} v_{kj}^* v_{mj} \sum_{s} v_{ks} v_{rs}^* \sum_{t} v_{lt} v_{mt}^* dt$$
$$= \sum_{s,t} \sum_{k} v_{kj}^* v_{ks} \sum_{m} v_{mj} v_{mt}^* \sum_{l,r} u_{li}^* u_{ri} v_{rs}^* v_{lt} dt$$
$$= \sum_{l,r} u_{li}^* u_{ri} v_{rj}^* v_{lj} dt = |\langle u_i(t), v_j(t)\rangle|^2 dt \ .$$

□

Lemma 5.4 tells us that the covariation between the driving Brownian motions of eigenvalues $\lambda \in sp(H)$ and $\mu \in sp(\widetilde{H})$ is given by the square of the inner product between their corresponding eigenvectors. Without loss of generality we may take H to be diagonal. Then $u_i = e_i$ for all $1 \le i \le n$, where $\{e_1, ..., e_n\}$ is the canonical basis for $\mathbb{C}^n$. Hence, to evaluate the bracket (5.13) we only need to find an expression for the eigenvectors of $\widetilde{H}$.

Consider $y = (y_1, ..., y_{n-1}) \in \mathbb{C}^{n-1}$ and let $x = (x_1, ..., x_n) = y_1 v_1 + ... + y_{n-1} v_{n-1} \in L$. We have, using $x_i = \sum_j y_j v_{ij}$,

$$x^* H x = \sum_i \lambda_i |x_i|^2 = \sum_j |y_j|^2 (\lambda_1 |v_{1j}|^2 + ... + \lambda_n |v_{nj}|^2) + \sum_{k \neq j} y_k^* y_j (\lambda_1 v_{1k}^* v_{1j} + ... + \lambda_n v_{nk}^* v_{nj}) \ . \tag{5.15}$$

At the other hand

$$x^* \widetilde{H} x = \langle \sum_j y_j v_j, \widetilde{H} \sum_j y_j v_j \rangle = \sum_j |y_j|^2 \mu_j \ . \tag{5.16}$$

But for any $x \in L$, $x^* H x = x^* \widetilde{H} x$, and so equating coefficients of y_j's in (5.15) and (5.16), we arrive at a system of quadratic equations in v

$$\begin{array}{ccccccc} \mu_1 & = & \lambda_1 |v_{11}|^2 + & ... & + \lambda_n |v_{n1}|^2 & = & \langle v_1, H v_1 \rangle \ , \\ \vdots & & \vdots & & \vdots & & \vdots \\ \mu_n & = & \lambda_1 |v_{1n}|^2 + & ... & + \lambda_n |v_{nn}|^2 & = & \langle v_n, H v_n \rangle \ , \\ 0 & = & \lambda_1 v_{1i}^* v_{1j} + & ... & + \lambda_n v_{ni}^* v_{nj} & = & \langle v_i, H v_j \rangle \ , \end{array} \tag{5.17}$$

for $i \neq j$.

Now write $Hv_j \in \mathbb{C}^n$, $1 \leq j \leq n$, in the new $\mathbb{C}^n$-basis $\{v_1, ..., v_{n-1}, l\}$ as follows

$$Hv_j = \sum_i^{n-1} \alpha_i v_i + \gamma_j l = \alpha_j v_j + \gamma l = \mu_j v_j + \gamma_j l \ ,$$

where $\alpha_1, ..., \alpha_{n-1}, \gamma_j \in \mathbb{C}$. The last two equality signs follow because by (5.17) we must have $v_i^* H v_j = 0$ for all $i \neq j$ and $v_j^* H v_j = \mu_j$. Hence, for $1 \leq i \leq n$ we have

$$\begin{aligned} & \lambda_i v_{ij} = \mu_j v_{ij} + \gamma_j l_i \\ \Leftrightarrow \qquad & v_{ij} = \frac{\gamma_j l_i}{\lambda_i - \mu_j} \ . \end{aligned}$$

Recall that our goal is to express eigenvectors v in terms of μ's and λ's. Note that γ_j is just a quantity ensuring that $\sum_i |v_{ij}|^2 = 1$. Hence, to accomplish our task we need to find an expression for l in terms of eigenvalues of H and $\widetilde{H}$.

For any $\mu \in sp(\widetilde{H})$ and the corresponding eigenvector $v^\star = (v_1^\star, ..., v_n^\star)$ we write $\widetilde{H} v^\star = Hv^\star - \langle Hv^\star, l \rangle l = \mu v^\star$, or $(H - \mu \mathbb{I}) v^\star - \langle Hv^\star, l \rangle l = 0$, where $\mathbb{I}$ is an identity matrix of an appropriate size. Together with $\langle l, v^\star \rangle = 0$, this gives us a system of equations

$$\begin{array}{ccccccc} (\lambda_1 - \mu) v_1^\star & +0 & \cdots & +0 & - & \langle Hv^\star, l \rangle l_1 & = & 0\,, \\ 0 & +(\lambda_2 - \mu) v_2^\star & \cdots & +0 & - & \langle Hv^\star, l \rangle l_2 & = & 0\,, \\ \vdots & \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & +0 & \cdots & +(\lambda_n - \mu) v_n^\star & - & \langle Hv^\star, l \rangle l_n & = & 0\,, \\ l_1^* v_1^\star & +l_2^* v_2^\star & \cdots & +l_n^* v_n^\star & & & = & 0\,. \end{array} \tag{5.18}$$

in $(v_1^\star, ..., v_n^\star)$.

We know that a determinant of a square matrix is zero if and only if its columns are not linearly independent. In other words if A is an $n \times n$ matrix, then $\det(A) = 0$ iff there exists a non-zero vector $\alpha = (\alpha_1, ..., \alpha_n)$ such that $\alpha_1 A_1 + ... + \alpha_n A_n = 0$, where A_i

is the i^{th} column of A. It follows that equations (5.18) hold true if and only if

$$\begin{vmatrix} \lambda_1-\mu & \cdots & 0 & l_1 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \lambda_n-\mu & l_n \\ l_1^* & \cdots & l_n^* & 0 \end{vmatrix} = 0 \, .$$

The above is equivalent to the condition

$$\prod_{i=1}^{n}(\lambda_i-\mu)\sum_{k=1}^{n}\frac{w_k}{\lambda_k-\mu}=0 \, ,$$

where $w_k = |l_k|^2$. Since the above equation must hold for all $\mu \in sp(\widetilde{H})$, we can find $\{w_1, ..., w_n\}$ by solving the following system of n equations

$$\begin{array}{ccccccc} \frac{w_1}{\lambda_1-\mu_1} & + & \cdots & + & \frac{w_n}{\lambda_n-\mu_1} & = & 0 \, , \\ \vdots & & & & \vdots & & \vdots \\ \frac{w_1}{\lambda_1-\mu_{n-1}} & + & \cdots & + & \frac{w_n}{\lambda_n-\mu_{n-1}} & = & 0 \, , \\ w_1 & + & \cdots & + & w_n & = & 1 \, , \end{array}$$

where the last equation comes from the fact that $|l|^2 = \sum_i w_i = 1$. To solve for w_i's we need

Lemma 5.5. *For $x = (x_1, ..., x_n) \in \mathbb{R}^n$ and $y = (y_1, ..., y_n) \in \mathbb{R}^n$ we have*

$$det\left(\frac{1}{x_i-y_j}\right)_{1\le i,j\le n} = (-1)^{\alpha}\frac{\triangle(x)\triangle(y)}{\prod_{1\le i,j\le n}(x_i-y_j)} \, , \tag{5.19}$$

where $\alpha = n/2$ if n is even and $\alpha = (n-1)/2$ if n is odd.

Proof See appendix at the end of the chapter.

Let A be a $n \times n$ invertible matrix and $x = (x_1, ..., x_n)$ and $y = (y_1, ..., y_n)$ two vectors such that $Ax = y$. Then the Cramer's rule states that $x_m = \frac{\det A[m]}{\det A}$, for $1 \le m \le n$, where $A[m]$ is a matrix obtained from A by substituting m^{th} column of A

with vector y. Write the above system of equations in the matrix form

$$\begin{pmatrix} \frac{1}{\lambda_1-\mu_1} & \cdots & \frac{1}{\lambda_n-\mu_1} \\ \vdots & \ddots & \vdots \\ \frac{1}{\lambda_1-\mu_{n-1}} & \cdots & \frac{1}{\lambda_n-\mu_{n-1}} \\ 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_{n-1} \\ w_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} .$$

Then, coupled with the identity of Lemma 5.5, Cramer's rule gives us

$$w_i = \frac{\triangle(\lambda^{(i)}) \prod_{1\leq k\leq n-1}(\lambda_i - \mu_k)(-1)^i}{\sum_{m=1}^n \triangle(\lambda^{(m)}) \prod_{1\leq k\leq n-1}(\lambda_m - \mu_k)(-1)^m} ,$$

where, for all $1 \leq m \leq n$, we write $\lambda^{(m)}$ for the vector λ with m^{th} element deleted. Now, consider a polynomial in λ's and μ's comprising the numerator of the rational function above. Suppose $\lambda_p = \lambda_l$ for any two indices $p \neq l$. Then

$$\begin{aligned} &\sum_{m=1}^n \triangle(\lambda^{(m)}) \prod_{1\leq m\leq n-1} (\lambda_m - \mu_k)(-1)^k \\ &\quad = \triangle(\lambda^{(p)}) \prod_{1\leq p\leq n-1} (\lambda_p - \mu_k)(-1)^p + \triangle(\lambda^{(l)}) \prod_{1\leq k\leq n-1} (\lambda_p - \mu_k)(-1)^p \\ &\qquad = \prod_{1\leq k\leq n} (\lambda_p - \mu_k)[\triangle(\lambda^{(p)})(-1)^p - \triangle(\lambda^{(l)})(-1)^l] = 0 , \end{aligned}$$

which means that the numerator can be written as a product $\prod_{i<j}(\lambda_i - \lambda_j) = \triangle(\lambda)$ times some constant; this constant is in fact "-1" as one might find out by calculating the sign of the leading term of the polynomial and using the definition of $\triangle(\lambda)$. It follows that

$$w_i = \frac{\triangle(\lambda^{(i)}) \prod_{1\leq k\leq n-1}(\lambda_i - \mu_k)(-1)^i}{-\triangle(\lambda)} = \frac{\prod_{1\leq k\leq n-1}(\lambda_i - \mu_k)}{\prod_{\substack{1\leq k\leq n \\ k\neq i}}(\lambda_i - \lambda_k)} .$$

This finally gives us an expression for the eigenvector $v_j = (v_{1j}, ..., v_{nj})$ of μ_j, $1 \leq j \leq n-1$,

$$|v_{ij}|^2 = \frac{\gamma_j^2}{(\lambda_i - \mu_j)} \frac{\prod_{\substack{1\leq k\leq n-1 \\ k\neq j}}(\lambda_i - \mu_k)}{\prod_{\substack{1\leq k\leq n \\ k\neq i}}(\lambda_i - \lambda_k)} ,$$

and so by Lemma 5.4 $d\langle \alpha_i(t), \beta_j(t)\rangle = |v_{ij}(t)|^2$, where α_i and β_j are Brownian motions driving λ_i and μ_j respectively. Finally, because v_j is a unit vector, we have $\sum_i |v_{ij}|^2 = 1$, which gives the required expression (5.10) for γ_j^2. This completes the proof.

□

Remark. By calculating the distribution of the random vector w and finding the Jacobian of the transformation associating w's to eigenvalues μ, Baryshnikov shows that the probability density of the vector of ordered eigenvalues μ of the restriction $\widetilde{H}$ is proportional to the Vandermonde determinant $\triangle(\mu)$ (see [6, Prop. 4.2]). In effect we can generate a random GC pattern with the distribution of the Hermitian minor process at any $t > 0$ by first considering $H(t)$ and then its restrictions to a nested sequence of spaces $L_1 \subset L_2 \subset \ldots \subset L_n = \mathbb{C}^n$, with $L_k \simeq \mathbb{C}^k$, chosen from all the possible such filtrations under an $\mathscr{U}_n$-invariant distribution. Interlacing of the corresponding eigenvalues is again ensured by the Cauchy's interlacing theorem.

To illustrate the above approach, consider the easiest $n = 2$ case, i.e. the triangle array

$$\begin{array}{ccc} \lambda_1(t) & & \lambda_2(t) \\ & \mu(t) & \end{array} ,$$

where $(\lambda_1(t), \lambda_2(t))$ are eigenvalues of the Hermitian form

$$\begin{pmatrix} h_{11}(t) & \frac{h_{12}(t)+ih_{21}(t)}{\sqrt{2}} \\ \frac{h_12(t)-ih_{21}(t)}{\sqrt{2}} & h_{22}(t) \end{pmatrix}$$

and μ is the eigenvalue of its restriction to a random hyperplane $L \simeq \mathbb{C}$. Let $v(t) = (v_1(t), v_2(t))$ be the normalised eigenvector of $\mu(t)$. Equations (5.17) then translate simply to

$$\lambda_1|v_1|^2 + \lambda_2|v_2|^2 = \mu \ .$$

In this special case v_i's can be simply solved for by using $|v_1|^2 + |v_2|^2 = 1$. We obtain

$$|v_1|^2 = \frac{\mu - \lambda_2}{\lambda_1 - \lambda_2} \quad \text{and} \quad |v_2|^2 = \frac{\lambda_1 - \mu}{\lambda_1 - \lambda_2} ,$$

which coincides with the results obtained in example 3.4 in Chapter 3 by simply using Itô's formula.

5.2 Multidimensional Pitman's theorem and non-colliding processes

In this section we discuss the extension of the Pitman's theorem, both its classical and discrete versions, to higher dimensions. In notation of Chapter 3 this is the $\theta=\infty$ and $q=0$ cases respectively. We start by describing the discrete setting, from which the multidimensional extension of the Pitman's theorem is obtained by taking the scaling limit.

Let $X=(X(k);k\geq 1)$ be an n-dimensional Markov chain taking values in $\mathbb{N}^n=\{0,1,...\}^n$. Denote by o the origin of $\mathbb{R}^n$ and by $b=\{e_1,...,e_n\}$ the canonical orthonormal basis of $\mathbb{C}^n$. Suppose $X(0)=o$ and the transition matrix is given by

$$P(x,x+e_i)=\frac{1}{n},\quad 1\leq i\leq n\,. \tag{5.20}$$

We call X the *multinomial random walk*.

Define the *discrete Weyl chamber* as the subset of $\mathbb{N}^n$

$$\mathscr{W}_n=\{(x_1,...,x_n)\in\mathbb{N}^n : x_1\geq x_2\geq\cdots\geq x_n\}$$

and a function $h:\mathscr{W}_n\to\mathbb{R}^+$ by

$$h(x)=\prod_{i<j}\frac{\tilde{x}_i-\tilde{x}_j}{j-i}\ , \tag{5.21}$$

where $\tilde{x}_i=x_i-i$. We note that by definition $h(x)=|\mathscr{K}_x|$, where $\mathscr{K}_x$ is the collection of all Gelfand-Cetlin patterns with top row x, as defined in Chapter 2 (see Thm. 2.9).

Recall the branching rule for representations of $U(\mathfrak{gl}(n))$ (or $U_q(\mathfrak{gl}(n))$) (see Ch. 2.5). The dimensions of spaces on either side of the expression (2.16) are the same.

Using this and the fact that $dim(V_x) = |\mathcal{K}_x|$ we obtain the identity

$$n|\mathcal{K}_x| = \sum_{x:x\nearrow x'} |\mathcal{K}_{x'}| \,, \tag{5.22}$$

where $x \nearrow x'$ means that x is interlaced with x', i.e. $x \preceq x'$, and $|x|+1 = |x'|$. Define $\widehat{P}(x,y) = P(x,y)\,\mathbf{1}_{\{y\in\mathcal{W}_n\}}$ for $x-y \in b$. Using (5.22) we have

$$\sum_{i=1}^{n} \widehat{P}(x, x+e_i)h(x+e_i) = h(x), \quad x \in \mathcal{W}_n \,,$$

which shows that $h(x)$ is harmonic with respect to P restricted to $\mathcal{W}_n$. Thus

$$Q^{(n)}(x, x+e_i) := \widehat{P}(x, x+e_i)\frac{h(x+e_i)}{h(x)} = \frac{1}{n}\frac{|\mathcal{K}_{x+e_i}|}{|\mathcal{K}_x|}\mathbf{1}_{\{x+e_i\in\mathcal{W}_n\}}, \quad x \in \mathcal{W}_n \tag{5.23}$$

is a well defined transition matrix. The Markov process $\widehat{X}$ defined by these transition probabilities and started at o is a system of n symmetric random walks conditioned to stay in the Weyl chamber $\mathcal{W}_n$, i.e. to maintain their order. We will refer to this process as *Dyson's random walks*.

In [59] O'Connell gives a construction of Dyson's random walks by applying a certain transformation to the original process X. The construction produces a process in the Gelfand-Cetlin cone, such that the top row is evolving as the n-dimensional Dyson's random walk. This transformation is closely related to the Robinson-Schensted algorithm and also has an interpretation in queueing theory. We discuss it next.

5.2.1 A representation for conditioned random walks

In this section we describe certain transformations for random walks introduced by O'Connell and Yor [60]. Continuous version of these transformations can be seen as a generalisation of the Pitman's transform to the Weyl chambers and also have intimate connection to the eigenvalues of a GUE matrix.

Let $D_0(\mathbb{N})$ denote the set of paths $x : \mathbb{N} \to \mathbb{N}$ with $x(0) = 0$, and for $x, y \in D_0(\mathbb{N})$

define

$$(x \triangle y)(m) = \min_{0 \le k \le m} [x(k) + y(m) - y(k)] ,$$
$$(x \bigtriangledown y)(m) = \max_{0 \le k \le m} [x(k) + y(m) - y(k)] . \tag{5.24}$$

For $n \ge 2$ define the mappings $G^{(n)} : D_0(\mathbb{N})^n \to D_0(\mathbb{N})^n$ as follows. For $n = 2$ let

$$G^{(2)}(x, y) = (x \triangle y, y \bigtriangledown x) ,$$

and for $n \ge 2$ define inductively

$$G^{(n)}(x_1, ..., x_n) = (x_1 \triangle \cdots \triangle x_n, G^{(n-1)}(x_2 \bigtriangledown x_1, x_3 \bigtriangledown (x_1 \triangle x_2), ..., x_n \bigtriangledown (x_1 \triangle \cdots \triangle x_{i-1}))) .$$

One can show that for each $n \ge 2$ and $x \in \mathbb{N}^n$, $G_1^{(n)}(x) \le G_2^{(n)}(x) \le \cdots \le G_n^{(n)}(x)$ (see O'Connell [59, Sec. 2]) and that, in particular, $(G^{(n)}(x_1, ..., x_n)^*, ..., G^{(1)}(x_1)^*)$ is a Gelfand-Cetlin pattern of depth n. Here we write x^*, for $x \in \mathbb{R}^n$, to indicate $x_i = x^*_{n-i+1}$, $1 \le i \le n$. It follows that if X is a random walk in $\mathbb{N}^n$ defined at the beginning of this section, then the transformed process $(G^{(i)}(X_1(k), ..., X_i(k))^*, 1 \le i \le n; k \ge 0)$ is a random walk in the Gelfand-Cetlin cone $\mathscr{K}_n$. Moreover,

Theorem 5.6. *([59, Cor. 6.2]) The transformed process $G^{(n)}(X)$ is distributed as the random walk X conditioned to stay in the Weyl chamber $\mathscr{W}_n$. In particular, the transition matrix of $G^{(n)}(X)$ is given by* (5.23).

Note that in case $n = 2$, Theorem 5.6 is equivalent to the discrete Pitman's theorem. Indeed, if we let $z = y - x$ and $m \in \mathbb{N}$, then

$$G_2^{(2)}(x, y)(m) - G_1^{(2)}(x, y)(m) = (y \bigtriangledown x)(m) - (x \triangle y)(m) = 2 \max_{1 \le k \le m} z(k) - z(m) .$$

But, since the two components of $G^{(2)}(X, Y)$ are conditioned to stay ordered, $G_2^{(2)}(X, Y) - G_1^{(2)}(X, Y)$ is distributed as a random walk conditioned to stay non-negative, i.e. as a discrete BES^3 process. Finally note that $Z = Y - X$ is just a simple symmetric random walk.

Just as in the classical Pitman's theorem, intertwining plays an important role; transition functions of the transformed and the original processes are intertwined with respect to a certain Markov kernel (see [59, Cor. 6.5]).

The connection to the Robinson-Schensted algorithm is as follows. Consider a word $a \in [n]^k$ and let $x_i(m) = |\{1 \le j \le m : a_j = i\}|$, for $1 \le i \le n$ and $1 \le m \le k$, i.e. $x_i(m)$ is the number of a_j's, up to and including a_m, equal to i. Denote by $\tau(m)$ the semi-standard Young tableau obtained by applying the Robinson-Schensted algorithm with column insertion to $(a_1, ..., a_m)$. Then for each $1 \le m \le k$, the GC pattern $((G^{(n)}(x_1, ..., x_n)^*)(m), ..., (G^{(1)}(x_1)^*)(m))$ corresponds to the semi-standard Young tableau $\tau(m)$ (see [59, Thm. 3.1]).

Now consider a sequence of random variables $(\eta_k, k \ge 0)$ such that $\eta_0 = 0$ and for each $i \ge 1$ η_i is distributed uniformly over $\{1, ..., n\}$. Define

$$X_i(m) = |\{1 \le j \le m : \eta_j = i\}|, \quad 1 \le i \le n, m \ge 0 .$$

Then $X = (X_1, ..., X_n)$ is the familiar multinomial random walk. Moreover, at any time $k \ge 1$, $X(k)$ is the type and $G^{(n)}(X(k))^*$ is the shape of a randomly growing Young tableau constructed from a random word $(\eta_1, ..., \eta_k)$ via column insertion. For two alternative dynamics related to the Robinson-Schensted-Knuth algorithm see [58].

In order to pass to the continuous version of the Theorem 5.6 and, thus, to a result about Hermitian Brownian motion, we define a Poissonized version of the Robinson-Schensted dynamics described above. Transformations (5.24), and so mappings $G^{(i)}$, admit continuous versions. Denote by $D_0(\mathbb{R}^+)$ a space of càdlàg paths $f : \mathbb{R}^+ \to \mathbb{R}$ with $f(0) = 0$ and let

$$(f \triangle g)(t) = \inf_{0 \le s \le t} [f(s) + g(t) - g(s)] ,$$

$$(f \bigtriangledown g)(t) = \sup_{0 \le s \le t} [f(s) + g(t) - g(s)] ,$$

for $f, g \in D_0(\mathbb{R}^+)$. Define

$$\Gamma^{(2)}(f, g) = (f \triangle g, g \bigtriangledown f)$$

and mappings $\Gamma^{(n)} : D_0(\mathbb{R}^+)^n \to D_0(\mathbb{R}^+)$, for $n \geq 2$, by

$$\Gamma^{(n)}(f_1, ..., f_n) = (f_1 \triangle \cdots \triangle f_n, \Gamma^{(k-1)}(f_2 \bigtriangledown f_1, f_3 \bigtriangledown (f_1 \triangle f_2), ..., f_k \bigtriangledown (f_1 \triangle \cdots \triangle f_{n-1}))) .$$

Let $N = (N_1, ..., N_n)$ with $N(0) = o$ be the counting processes of n independent Poisson processes on $\mathbb{R}^+$, each with intensity 1. Let $\widehat{N}$ be the h-transform of N with respect to $h(x)$ given by (5.21). Then the law of $\Gamma^{(n)}(N)$ is the same as the law of $\widehat{N}$ ([59, Thm. 7.1], [60, Thm. 5]). As in the discrete case, generators of the two processes are intertwined with respect to a certain Markov kernel (see [59, Thm. 7.2]).

Both the discrete conditioned random walk and its Poissonised version are closely connected to discrete orthogonal polynomial ensembles and so to determinantal processes. In particular, the distribution of the Poisson random walk is connected to the *Charlier ensemble*, while the distribution of the multinomial walk – to the *de-Poissonised Charlier ensemble*, see [54] and also [44]. It is this determinantal structure of the measures involved that lies at the heart of the relation of both processes to the GUE matrices.

By applying an appropriate version of the Dönsker's theorem one obtains a version of Theorem 5.6 for the Brownian motion. Let B be a standard n-dimensional Brownian motion started at the origin and $\hat{B}$ an h-transform of B with $h(x)$ given by the Vandermonde determinant. Then

Theorem 5.7. *(O'Connell, Yor [60, Thm. 7]) The processes $\hat{B}$ and $\Gamma^{(n)}(B)$ have the same law.*

Again when $n = 2$, this result corresponds to the Pitman's construction of the 3-dimensional Bessel process. Since by Theorem 5.7 $\Gamma^{(2)}(B(t))$ is a 2-dimensional Dyson's BM, one verifies with the help of Itô's lemma that $R(t) := (\Gamma^{(2)}_2(B(t)) - \Gamma^{(2)}_1(B(t)))/\sqrt{2}$ has the distribution of the BES^3 process. At the same time $X(t) := (\Gamma^{(2)}_2(B(t)) + \Gamma^{(2)}_1(B(t)))/\sqrt{2}$ is a standard Brownian motion independent of R. But by definition of $\Gamma^{(1)}$

$$R(t) = \frac{1}{\sqrt{2}}(\Gamma^{(2)}_2(B(t)) - \Gamma^{(2)}_1(B(t))) = 2 \sup_{s \leq t} X(s) - X(t) .$$

Recalling that the eigenvalues of an $n \times n$ Hermitian BM are distributed as n

Brownian motions conditioned to stay in the Weyl chamber, we see that Theorem 5.7 yields an alternative representation for them. In particular, by showing that $\Gamma_n^{(n)}(B) = (B_1 \triangledown \cdots \triangledown B_n)$, one finds an alternative proof of the result of Baryshnikov [6] and Tracy, Gravner and Widom [37] which states that the random variable

$$\lambda_1 := \Gamma_n^{(n)}(B(1)) = \sup_{0 \le t_1 \le \ldots \le t_n = 1} \sum_{k=1}^{n} \{B_k(t_k) - B_k(t_{k-1})\}$$

has the same distribution as the largest eigenvalue of a GUE matrix. What's more, we see that O'Connell and Yor's results generalise this formula to give a description of all the eigenvalues of a GUE matrix.

The same transformations were also described by Bougerol and Jeulin [14] in a purely representation theoretic context. Authors describe a transformation taking paths in a finite-dimensional Euclidian space $\mathfrak{a}$ to paths in the interior of the corresponding Weyl chamber $\mathfrak{a}_+$. $\mathfrak{a}$ is taken to be the maximal Torus of the maximal compact subalgebra of some algebra $\mathfrak{g}$. In case of a Hermitian Brownian motion $\mathfrak{g} = \mathfrak{gl}_{\mathbb{C}}(n)$ and $\mathfrak{a} \simeq \mathbb{R}^n$ is the space of all $n \times n$ real diagonal matrices. By taking $\mathfrak{g}$ to be other classical groups Bougerol and Jeulin find representation for eigenvalues of other types of random matrices. See also a paper by Biane, Bougerol and O'Connell [12].

5.3 A Markov chain in the Gelfand-Cetlin cone

In this section we construct and analyse an higher-dimensional generalisation of the Markov chain $(\mathrm{RW}(\mu), \mathrm{dBES}^3(\mu))_q$ of Chapter 3. We will identify several intertwining relationships and present an informal argument relating a certain Markov function of this process, when $q = 1$, to the Hermitian minor process.

We start by recalling some notation that will be heavily used in this section. A Gelfand-Cetlin cone is a collection of all integer Gelfan-Cetlin patterns of depth n:

$$\mathscr{K}_n = \{(m^n, m^{n-1} \ldots, m^1) \in \mathbb{N}^n \times \mathbb{N}^{n-1} \times \ldots \times \mathbb{N} : m^n \succeq m^{n-1} \succeq \ldots \succeq m^1\}\ .$$

By $\mathscr{K}_{(z^n,\ldots,z^p)}$, with $z^n \succeq \ldots \succeq z^p$, we denote a subset of $\mathscr{K}_n$ with top $n-p+1$

rows given by $(z^n,...,z^p)$. In particular $\mathcal{K}_\lambda \subset \mathcal{K}_n$ denotes a subset of all patterns of depth n with top row λ and $\mathcal{K}_{(\lambda,\mu)} \subset \mathcal{K}_\lambda \subset \mathcal{K}_n$ denotes a subset of all patterns of depth n with top row λ and second row μ. By $(z^n,...,z^p,\xi) \in \mathcal{K}_{(z^n,...,z^p)}$ we denote a pattern with top $n-p+1$ rows given by $(z^n,...,z^p)$ and the 'rest of the pattern' given by ξ; by $(z^n,...,z^1) \in \mathcal{K}_n$ we mean a pattern such that for all $1 \le m \le n$ the m^{th} row is given by z^m. From Chapter 2 we know that for each Young diagram of shape $\lambda \in \mathcal{W}_n$ there is an irreducible representation V_λ of $U_q(\mathfrak{gl}(n))$ (and $U(\mathfrak{gl}(n))$) of dimension $|\mathcal{K}_\lambda|$ spanned by vectors $\{e_m; m \in \mathcal{K}_\lambda\}$. By e^λ_ξ, resp. $e^\lambda_{(\mu,\xi)}$, we will mean a basis vector of V_λ corresponding to a pattern $(\lambda,\xi) \in \mathcal{K}_\lambda$, resp. $(\lambda,\mu,\xi) \in \mathcal{K}_{(\lambda,\mu)}$. Finally, as before, we will write $\lambda \nearrow \lambda'$ to say that $\lambda \preceq \lambda'$ and $|\lambda|+1 = |\lambda'|$; $\{e_1,...,e_n\}$ denotes the canonical basis of $\mathbb{C}^n$.

We consider a discrete-time Markov chain $\mathcal{Z}_q := (\mathcal{Z}(k); k \ge 1) = (\mathcal{Z}^i(k); 1 \le i \le n, k \ge 1)$ defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P}^{(n)}_q)$ and taking values in $\mathcal{K}_n$. Let $y = (y^n,...,y^1), z = (z^n,...,z^1) \in \mathcal{K}_n$ be such that $z^m = y^m + e_{i_m}$ for $1 \le j \le m \le n$ and $z^m = y^m$ otherwise; we write $y = (z; i_n,...,i_j)$. We characterise the process $\mathcal{Z}$ by defining for any such patterns y and z the one-step transition probabilities, for $k \ge 0$, to be

$$\mathbb{P}^{(n)}_q(\mathcal{Z}(k+1) = z | \mathcal{Z}(k) = y) = P^{(n)}_q(y,z) = \frac{1}{n} w_q^2((z; i_n,...,i_j)) = \frac{1}{n}\sum_{k=1}^{n} |\langle e_y \otimes e_k, e'_z\rangle|^2$$

and 0 otherwise. Here $e_y \in V_{y^n}$, $e'_z \in V_{z^n} \subset V_{y^n} \otimes \mathbb{C}^n$ and $w_q((z; i_n,...,i_j))$ is a $U_q(\mathfrak{gl}(n))$-Wigner coefficient (see Ch. 2, page 44). Thus, in the rightmost summation only one inner product is non-zero, namely, $|\langle e_y \otimes e_j, e'_z\rangle|$. When $q = 1$ we are in the 'classical' set-up corresponding to $U(\mathfrak{gl}(n))$. For all $q \in (0,1]$ the dynamics of the chain is such that at each time step exactly one particle in the top $i_j \in \{1,...,n\}$ rows makes a jump of size 1, i.e. for all $1 \le m \le n-1$, $\mathcal{Z}^m(k+1) - \mathcal{Z}^m(k) \in \{0, e_1,...,e_n\}$ and $\mathcal{Z}^n(k+1) - \mathcal{Z}^n(k) \in \{e_1,...,e_n\}$, $k \ge 0$. Note that one of the particles in the top row jumps at *each* time step. Definition of transition probabilities as squares of Wigner coefficients ensures that the process never leaves the space $\mathcal{K}_n$.

We point out that, by construction, the family of processes $\mathcal{Z}_q$ is a higher-dimensional generalisation of the family of bivariate Markov chains of Chapter 3. In particular if we let $n = 2$ and use change of variables $\mathscr{Z}_1^2 - \mathscr{Z}_2^2 = R$ and $2\mathscr{Z}_1^1 - \mathscr{Z}_1^2 - \mathscr{Z}_2^2 = X$ we will recover familiar transition probabilities for the pair $(X,R)_q$, $q \in (0,1]$. Observe that the process associated to $U_q(\mathfrak{sl}(2))$ is two-dimensional, as opposed to the three-point process associated to $U_q(\mathfrak{gl}(2))$. This is explained by the fact that any two representations $V_{(a_1,a_2)}$ and $V_{(b_1,b_2)}$ of $U_q(\mathfrak{gl}(2))$, associated to Young tableaux of shapes (a_1, a_2) and (b_1, b_2), respectively, are isomorphic representations of $\mathfrak{sl}(2)$ if and only if $a_1 - a_2 = b_1 - b_2 = c$, where c is some integer. This makes the last row redundant and explains the new variable $r = a_1 - a_2$.

Firstly we state and prove

Proposition 5.8. *Let $\mathscr{Z}(0)$ be distributed uniformly over $\mathscr{K}_{\lambda^*}$ for some $\lambda^* \in \mathscr{W}_n$, i.e. $\mathbb{P}_q^{(n)}(\mathscr{Z}(0) = z) = \frac{1}{|\mathscr{K}_{\lambda^*}|}\mathbf{1}_{\{z \in \mathscr{K}_{\lambda^*}\}}$. Then the top row of the array $(\mathscr{Z}^n(k); k \geq 0)$ is a $\mathscr{W}_n$-valued Markov process with the initial state λ^* and transition functions $Q^{(n)}(\cdot,\cdot)$ given by* (5.23).

We remark that, in view of the connection of the representations of $U_q(\mathfrak{gl}(n))$ to the Robinson-Schensted algorithm in the limit $q \to 0$ (see end of Ch. 2), this proposition can be seen as a generalisation of Theorem 5.6.

See a paper by Warren and Windridge [80] for examples of alternative dynamics in the Gelfand-Cetlin cone, both in discrete and continuous time, and some related intertwining relationships.

Proof. We prove the proposition by identifying an intertwining relationship between the top row process $\mathscr{Z}^n$ and the whole pattern $\mathscr{Z}$ and appealing to Theorem 3.6. Consider a Markov kernel

$$\mathrm{M}_n(\lambda, z) = \frac{1}{|\mathscr{K}_\lambda|}\mathbf{1}_{\{z \in \mathscr{K}_\lambda\}}$$

from $\mathscr{W}_n$ to $\mathscr{K}_n$. Note that for each $\lambda \in \mathscr{W}_n$ $\mathrm{M}(\lambda,\cdot)$ constitutes the uniform law on the patterns with top row λ. Using the fact that $M_n(\lambda, z)$ is non-zero if and only if $z^n = \lambda$, one easily verifies that the intertwining relationship we would like to prove

$$\sum_{(\lambda'',\xi'')\in\mathcal{K}_n} \mathrm{M}_n(\lambda,(\lambda'',\xi''))P_q^{(n)}((\lambda'',\xi''),(\lambda',\xi')) = \sum_{\lambda''\in\mathcal{W}_n} Q^{(n)}(\lambda,\lambda'')\mathrm{M}_n(\lambda'',(\lambda',\xi'))$$

is equivalent to

$$\sum_{\xi:(\lambda,\xi)\in\mathcal{K}_\lambda} \frac{1}{|\mathcal{K}_\lambda|}P_q^{(n)}((\lambda,\xi),(\lambda',\xi')) = Q^{(n)}(\lambda,\lambda')\frac{1}{|\mathcal{K}_{\lambda'}|} \tag{5.25}$$

for all $\lambda \nearrow \lambda'$.

For the left-hand side we have, for any admissible choices of (λ,ξ) and (λ',ξ')

$$\sum_{\xi:(\lambda,\xi)\in\mathcal{K}_\lambda} \frac{1}{|\mathcal{K}_\lambda|}P_q^{(n)}((\lambda,\xi),(\lambda',\xi')) = \frac{1}{|\mathcal{K}_\lambda|}\sum_{\xi:(\lambda,\xi)\in\mathcal{K}_\lambda}\sum_{i=1}^{n}\frac{1}{n}|\langle e_\xi^\lambda \otimes e_i, e_{\xi'}^{\lambda'}\rangle|^2 = \frac{1}{n}\frac{1}{|\mathcal{K}_\lambda|}\ . \tag{5.26}$$

To see where the last equality comes from, consider two finite-dimensional vector spaces W and V with $W \subset V$. If $\{v_1,\dots,v_n\}$ and $\{w_1,\dots,w_k\}$ are orthonormal bases of V and W respectively, then one easily shows that $\sum_{i=1}^{n}|\langle v_i,w_j\rangle|^2 = 1$ for all $1\le j\le k$. We know that $\{e_\xi^\lambda\otimes e_i;(\lambda,\xi)\in\mathcal{K}_\lambda, 1\le i\le n\}$ and $\{e_{\xi'}^{\lambda'};(\lambda',\xi')\in\mathcal{K}_{\lambda'}\}$ are orthonormal bases of $U_q(\mathfrak{gl}(n))$-invariant spaces $V_\lambda\otimes\mathbb{C}^n$ and $V_{\lambda'}\subset V_\lambda\otimes\mathbb{C}^n$ respectively. It follows that $\sum_\xi\sum_{i=1}^{n}|\langle e_\xi^\lambda\otimes e_i, e_{\xi'}^{\lambda'}\rangle|^2 = 1$ which establishes (5.26).

On the other hand, for any $\lambda,\lambda'\in\mathcal{W}_n$ we have

$$Q^{(n)}(\lambda,\lambda')\frac{1}{|\mathcal{K}_{\lambda'}|} = \frac{1}{n}\frac{1}{|\mathcal{K}_\lambda|}, \quad \text{if } \lambda\nearrow\lambda'$$

and 0 otherwise, which establishes identity (5.25).

Evidently, all the other conditions of Theorem 3.6 are satisfied with $\phi(z) = \phi(z^n,z^{n-1}\dots,z^1) = z^n$, and so it follows that $\mathcal{Z}^n$ is a Markov chain with $\mathcal{Z}^n(0)=\lambda^*$ and the transition function $Q^{(n)}$. Moreover,

$$\mathbb{P}_q^{(n)}(\mathcal{Z}(k)=z|\mathcal{Z}^n(k)=\lambda,\mathcal{Z}^n(i)=y_i, 1\le i\le k-1) = \\ = \mathbb{P}_q^{(n)}(\mathcal{Z}(k)=z|\mathcal{Z}^n(k)=\lambda) = \frac{1}{|\mathcal{K}_\lambda|}\mathbf{1}_{\{z\in\mathcal{K}_\lambda\}}\ .$$

In other words, at any time $k \geq 1$, given the value of $\mathscr{Z}^n(k)$, the rest of the pattern is distributed uniformly over all admissible patterns with top row $\mathscr{Z}^n(k)$. This property is analogous to the result of Baryshnikov relating the distribution of the eigenvalues of the principal minors of a GUE matrix given the eigenvalues of the matrix itself.

□

Just like in the case $n = 2$ of Chapter 3, the marginals of the process $(X,R)_q$ are distributed as a SSRW and a discrete BES(3) process irrespective of the value of the parameter q, the marginal process $\mathscr{Z}^n$ is distributed as Dyson random walk for all $q \in (0,1]$. Recall that in the case $q = 1$ R was a Markov process through the Dynkin criteria as well as by the intertwining condition. Below is the n-dimensional version of that result.

Proposition 5.9. *In the case $q = 1$ we have*

$$\mathbb{P}^{(n)}(\mathscr{Z}^n(k) = \lambda | \mathscr{Z}(k-1) = z) = \mathbb{P}^{(n)}(\mathscr{Z}^n(k) = \lambda | \mathscr{Z}^n(k-1) = z^n)$$

for all $z \in \mathscr{K}_n$ with $z^n \nearrow \lambda$, $k \geq 1$. In particular,

$$\sum_{\xi'} P^{(n)}((\lambda, \xi), (\lambda', \xi')) = Q^{(n)}(\lambda, \lambda') \tag{5.27}$$

for all admissible $(\lambda, \xi), (\lambda', \xi') \in \mathscr{K}_n$.

The proof requires a simple result from linear algebra.

Lemma 5.10. *Consider a finite-dimensional vector space V and subspaces $U, W \subseteq V$ with orthonormal bases $\{u_1, ..., u_p\}$ and $\{w_1, ..., w_l\}$ respectively. Then*

$$\sum_{i,j} |\langle u_i, w_j \rangle|^2 = tr(P_U P_W) , \tag{5.28}$$

where P_U and P_W are orthogonal projections on U and W respectively. In particular, the LHS of (5.28) *doesn't depend on the choice of bases for U and W.*

Proof. Denote an orthonormal basis of V by $\{v_1, ..., v_n\}$, and suppose that the embeddings of U and W in V are given by $u_i = \sum_{k=1}^n \alpha_k^i v_k$ and $w_i = \sum_{k=1}^n \beta_k^i v_k$ respectively, with $\beta_k \in \mathbb{C}$, $\alpha_k \in \mathbb{C}$ for $1 \le k \le n$. Then one computes the LHS of (5.28) to be equal to $\sum_{i,j} |\langle \alpha^i, \beta^j \rangle|^2$, where $\alpha^i = (\alpha_1^i, ..., \alpha_n^i)$, $1 \le i \le p$, and $\beta^j = (\beta_1^j, ..., \beta_n^j)$, $1 \le j \le l$. On the other hand

$$\begin{aligned} tr(P_U P_W) = tr\left(\sum_i u_i u_i^* \sum_j w_j w_j^*\right) &= tr\left(\sum_{i,j} u_i w_j^* \langle u_i, w_j\rangle\right) \\ &= \sum_{i,j} \langle \alpha^i, \beta^j \rangle tr\left(u_i w_j^*\right) = \sum_{i,j} |\langle \alpha^i, \beta^j\rangle|^2 \ , \end{aligned}$$

where the last equality follows from the fact that $tr(u_i w_j^*) = \sum_{m,k} \alpha_k^i \beta_m^{*j} \langle v_k, v_m \rangle$. Note that for the special case when $U = V$ we have

$$tr(P_V P_W) = \sum_{i,j} |\langle v_i, w_j \rangle|^2 = dim(W) \ . \tag{5.29}$$

□

Proof of Proposition 5.9. For any $(\lambda, \xi) \in \mathscr{K}_\lambda$ and $\lambda \nearrow \lambda'$ one has

$$\begin{aligned} P^{(n)}((\lambda,\xi),\lambda') &= \sum_{\xi'} P^{(n)}((\lambda,\xi),(\lambda',\xi')) \\ &= \frac{1}{n}\sum_{\xi'}\sum_i |\langle e_\xi^\lambda \otimes e_i, e_{\xi'}^{\lambda'}\rangle|^2 \\ &= \frac{1}{n}\sum_{\xi'}\sum_i |\langle g(e_\xi^\lambda \otimes e_i), g(e_{\xi'}^{\lambda'})\rangle|^2 \qquad \text{for any } g \in SU(n) \\ &= \frac{1}{n}\sum_{\xi'}\sum_i |\langle g(e_\xi^\lambda) \otimes g(e_i), g(e_{\xi'}^{\lambda'})\rangle|^2 \\ &= \frac{1}{n}\sum_{\xi'}\sum_i |\langle g(e_\xi^\lambda) \otimes e_i, e_{\xi'}^{\lambda'}\rangle|^2 \ . \end{aligned}$$

We used Lemma 2.10 from Ch. 2 which states that the action of $g \in SU(n)$ on representation spaces V_λ, $V_{\lambda'}$ is unitary to go from line 2 to line 3. The equality between line 4 and the last line is justified by Lemma 5.10, once we notice that $\{g(e_\xi^\lambda) \otimes g(e_i), 1 \le$

$i \leq n\}$ and $\{g(e^{\lambda'}_{\xi'}) : \xi' : (\lambda', \xi') \in \mathscr{K}_{\lambda'}\}$ form normalised bases of $g(e^{\lambda}_{\xi}) \otimes \mathbb{C}^n$ and $V_{\lambda'}$ respectively, both being subspaces of $V_\lambda \otimes \mathbb{C}^n$. But since $g(e^{\lambda}_{\xi}) \in V_\lambda$ and the choice of $g \in SU(n)$ was arbitrary, $P((\lambda, \xi), \lambda')$ only depends on λ and not on the rest of the pattern ξ, which is what we needed to prove.

□

It turns out that the process formed by the top row $\mathscr{Z}^n$ is not the only Markov function of $\mathscr{Z}$, and that, in fact, the top $1 \leq m \leq n$ rows together also evolve as a Markov chain with respect to their own filtration. The distributions of such processes, however, are not independent of the parameter q.

Theorem 5.11. *Let $q \in (0,1]$. Suppose $\mathscr{Z}(0)$ is distributed uniformly over $\mathscr{K}_{(z^n_*,...,z^p_*)}$, $1 \leq p \leq n$, the collection of all GC patterns with top $n-p+1 := m$ rows given by $(z^n_*, ..., z^p_*)$. Then the marginal process $(\mathscr{Z}^n(k), ..., \mathscr{Z}^p(k); k \geq 0)$ is a Markov chain with values in $(\mathscr{W}_n \times \mathscr{W}_{n-1} ... \times \mathscr{W}_p) \cap \mathscr{K}_n$ started at $(z^n_*, ..., z^p_*)$ and with one-step transition functions given by*

$$Q^{(n,m)}_q((z^n, ..., z^l, z^{l+1}, ..., z^p), (\hat{z}^n, ..., \hat{z}^l, z^{l+1}, ..., z^p)) =$$
$$= \frac{1}{n} w^{(1)}_q \left(\begin{array}{c|c} \hat{z}^l & i_l \\ z^{l-1} & \end{array} \right)^2 \prod_{k=l+1}^{n} w^{(2)}_q \left(\begin{array}{c|c} \hat{z}^k & i_k \\ \hat{z}^{k-1} & i_{k-1} \end{array} \right)^2$$

for $l \in \{n, ..., p-1\}$, and

$$Q^{(n,m)}_q((z^n, ..., z^p), (\hat{z}^n, ..., \hat{z}^p) = \frac{1}{n} \frac{|\mathscr{K}_{\hat{z}^p}|}{|\mathscr{K}_{z^p}|} \prod_{k=p+1}^{n} w^{(2)}_q \left(\begin{array}{c|c} \hat{z}^k & i_k \\ \hat{z}^{k-1} & i_{k-1} \end{array} \right)^2,$$

where $\hat{z}^k = z^k + e_{i_k}$, $1 \leq k \leq p$, and $w^{(1)}_q$ and $w^{(2)}_q$ are reduced Wigner coefficients (see p. 45 and p. 47).

Proof. Again the statement of the theorem is proved via identifying a suitable intertwining relationship and appealing to Theorem 3.6. Consider a Markov kernel

$$\mathrm{M}_{n,m}((z^n, ..., z^p), x) = \frac{1}{|\mathscr{K}_{(z^n,...,z^p)}|} \mathbf{1}_{\{x \in \mathscr{K}_{(z^n,...,z^p)}\}}$$

from $(\mathscr{W}_n \times \ldots \times \mathscr{W}_p) \cap \mathscr{K}_n$ to $\mathscr{K}_n$. For any interlaced m-tuple $z^n \succeq \ldots \succeq z^p$ $\mathrm{M}_{n,m}((z^n, \ldots, z^p), \cdot)$ is a uniform measure on the patterns $x \in \mathscr{K}_{(z^n, \ldots, z^p)}$. Using the definition of the kernel M, one easily checks that the desired intertwining relationship

$$\sum_{x \in \mathscr{K}_n} \mathrm{M}_{n,m}((z^n, \ldots, z^p), x) P_q^{(n)}(x, \hat{z}) \\ = \sum_{x^n \succeq \ldots \succeq x^p} Q_q^{(n,m)}((z^n, \ldots, z^p), (x^n, \ldots, x^p)) \mathrm{M}_{n,m}((x^n, \ldots, x^p), \hat{z})$$

is equivalent to the equality

$$\sum_{\xi} \frac{1}{|\mathscr{K}_{z^p}|} P_q^{(n)}((z^n, \ldots, z^p, \xi), (\hat{z}^n, \ldots, \hat{z}^p, \hat{\xi})) = Q_q^{(n,m)}((z^n, \ldots, z^p), (\hat{z}^n, \ldots, \hat{z}^p)) \frac{1}{|\mathscr{K}_{\hat{z}^p}|}, \tag{5.30}$$

where we have used the fact that $|\mathscr{K}_{(z^n, \ldots, z^p)}| = |\mathscr{K}_{z^p}|$ and we write $(z^n, \ldots, z^p, \xi)$ for the GC pattern with top $m = n - p + 1$ rows $(z^n, \ldots, z^p)$ and 'the rest of the pattern' given by ξ.

We treat the situation when particles jump in the top $l \leq m - 1 = n - p$ rows only separately from the situation when particles jump in in the first $l \geq m$ rows only. The former is trivial, since in that case transition function $P_q^{(n)}$ only depends on the first $l \leq m$ rows and not on the rest of the pattern; relationship (5.30) then follows directly by using the definition of $P_q^{(n)}$ and noticing that $z^p = \hat{z}^p$ when $l \leq n - p$.

Consider now the case $l \geq m$. Note that the multiplicative structure of the Wigner coefficients means that we can write

$$P_q^{(n)}(z^n, \ldots, z^p, \xi), (\hat{z}^n, \ldots, \hat{z}^p, \hat{\xi})) = \\ = \frac{p}{n} \prod_{k=p+1}^{n} w_q^{(2)} \left(\begin{array}{c|c} \hat{z}^k & i_k \\ \hat{z}^{k-1} & i_{k-1} \end{array} \right)^2 P_q^{(p)}((z^p, \xi), (\hat{z}^p, \hat{\xi})), \tag{5.31}$$

where it is possible that $\hat{\xi} = \xi$. Using this, the left-hand side of (5.30) reads

$$\sum_{\xi} \frac{1}{|\mathcal{K}_{z^p}|} P_q^{(n)}((z^n, \dots, z^p, \xi), (\hat{z}^n, \dots, \hat{z}^p, \hat{\xi})) =$$
$$= \frac{p}{n} \prod_{k=p+1}^{n} w_q^{(2)} \left(\begin{array}{c|c} \hat{z}^k & i_k \\ \hat{z}^{k-1} & i_{k-1} \end{array} \right)^2 \sum_{\xi} \frac{1}{|\mathcal{K}_{z^p}|} P_q^{(p)}((z^p, \xi), (\hat{z}^p, \hat{\xi}))$$
$$= \frac{p}{n} \prod_{k=p+1}^{n} w_q^{(2)} \left(\begin{array}{c|c} \hat{z}^k & i_k \\ \hat{z}^{k-1} & i_{k-1} \end{array} \right)^2 \frac{1}{|\mathcal{K}_{\hat{z}^p}|} Q^{(p)}(z^p, \hat{z}^p) \ ,$$

where the last equality is justified by the intertwining relationship (5.25) applied to the array formed by the last p rows of the pattern. Using $Q^{(p)}(z^p, \hat{z}^p) = \frac{1}{p} \frac{|\mathcal{K}_{\hat{z}^p}|}{|\mathcal{K}_{z^p}|}$, for any $z^p \nearrow \hat{z}^p$, gives (5.30). All the other conditions of Theorem 3.6 are easily verified to be satisfied with $\phi(z) = \phi(z^n, z^{n-1}, \dots, z^1) = (z^n, z^{n-1} \dots, z^p)$. Thus, $(\mathcal{Z}^n, \dots, \mathcal{Z}^p)$ is Markov with transition functions $Q^{(n,m)}$. $\square$

Again in the case $q = 1$ the Dynkin condition is satisfied, i.e. the process $(\mathcal{Z}^n, \dots, \mathcal{Z}^p)$ is Markov with respect to the filtration of the whole chain.

Proposition 5.12. *Let $q = 1$. Then for all $k \geq 1$*

$$\mathbb{P}^{(n)}(\mathcal{Z}^n(k) = \hat{z}^n, \dots, \mathcal{Z}^p(k) = \hat{z}^p | \mathcal{Z}^n(k-1) = z^n, \dots, \mathcal{Z}^1(k-1) = z^1)$$
$$= \mathbb{P}^{(n)}(\mathcal{Z}^n(k) = \hat{z}^n, \dots, \mathcal{Z}^p(k) = \hat{z}^p | \mathcal{Z}^n(k-1) = z^n, \dots, \mathcal{Z}^p(k-1) = z^p)$$

for any admissible $z = (z^n, \dots, z^1) \in \mathcal{K}_n$ and $(\hat{z}^n, \dots, \hat{z}^p) \in (\mathcal{W}_n \times \dots \times \mathcal{W}_p) \cap \mathcal{K}_n$. In particular,

$$\sum_{\xi} P^{(n)}((z^n, \dots, z^p, \xi), (\hat{z}^n, \dots, \hat{z}^p, \hat{\xi})) = Q^{(n,n-p+1)}((z^n, \dots, z^p), (\hat{z}^n, \dots, \hat{z}^p)) \qquad (5.32)$$

for all admissible $(z^n, \dots, z^p, \xi), (\hat{z}^n, \dots, \hat{z}^p, \hat{\xi}) \in \mathcal{K}_n$.

Proof. Using (5.31) we obtain

$$
\begin{aligned}
P^{(n)}((z^n,...,z^p,\xi),(\hat{z}^n,...,\hat{z}^p)) = \sum_{\hat{\xi}} P^{(n)}((z^n,...,z^p,\xi),(\hat{z}^n,...,\hat{z}^p,\hat{\xi})) = \\
= \frac{p}{n}\prod_{k=p+1}^{n} w^{(2)}\left(\begin{array}{c|c} \hat{z}^k & i_k \\ \hat{z}^{k-1} & i_{k-1}\end{array}\right)^2 \sum_{\hat{\xi}} P^{(p)}((z^p,\xi),(\hat{z}^p,\hat{\xi})) \\
= \frac{p}{n}\prod_{k=p+1}^{n} w^{(2)}\left(\begin{array}{c|c} \hat{z}^k & i_k \\ \hat{z}^{k-1} & i_{k-1}\end{array}\right)^2 Q^{(p)}(z^p,\hat{z}^p)\ ,
\end{aligned}
$$

where the third line follows by equation (5.27) of Lemma 5.9. The case when particles move in less than m top rows is trivial and follows from the definition of $P^{(n)}$. □

In what follows we will focus our attention on the process $(\mathscr{Z}^n(k),\mathscr{Z}^{n-1}(k);k\geq 0) := (\mathscr{X}(k),\mathscr{Y}(k);k\geq 0)$ generated by the top two rows of the pattern in the 'classical' set-up $q=1$. By Theorem 5.11 $(\mathscr{X},\mathscr{Y})$ is a Markov chain with values in $\mathscr{W}_{n,n-1} := (\mathscr{W}_n\times\mathscr{W}_{n-1})\cap\mathscr{K}_n$, and by Proposition 5.8 $\mathscr{X}$ is distributed as the Dyson's random walk, a system of n simple random walks in the space $\mathscr{W}_n$. The proposition below tells us that the marginal process $\mathscr{Y}$ is also Markov.

Proposition 5.13. *The marginal process $(\mathscr{Y}(k);k\geq 0)$ is Markov with values in $\mathscr{W}_{n-1}$ and with transition functions*

$$
A^{(n-1)}(\mu,\mu+e_j) = \frac{1}{n}\frac{|\mathscr{K}_{\mu+e_j}|}{|\mathscr{K}_\mu|}, \qquad 1\leq j\leq n-1\ ,
$$

and

$$
A^{(n-1)}(\mu,\mu) = \frac{1}{n}\ .
$$

Remark. One might also expect that, in fact, each row of $\mathscr{Z}$ viewed on its own is a Markov process evolving as a certain version of Dyson random walk. This conjecture, however, is yet to be proved.

Recall that for any $\mu\preceq\lambda\in\mathscr{W}_n$ we write V_λ^μ for a subspace of the $U(\mathfrak{gl}(n))$-irrep. V_λ invariant under the action of $U(\mathfrak{gl}(n-1))$ and such that $V_\lambda^\mu\simeq V_\mu$. We will require the following lemma for the proof of the proposition.

Lemma 5.14. *For any $(\lambda, \mu) \in \mathscr{W}_{n,n-1}$ and μ' such that $\mu \nearrow \mu'$ we have*

$$\sum_{\lambda'} tr(P_{V_{\lambda'}^{\mu'}} P_{V_\lambda^\mu \otimes \mathbb{C}^{n-1}}) = |\mathscr{K}_{\mu'}| \,, \tag{5.33}$$

where the sum is over all partitions λ' such that $\lambda \nearrow \lambda'$ and $\mu' \preceq \lambda'$, and projections $P_{V_{\lambda'}^{\mu'}}, P_{V_\lambda^\mu}$ are operators on the complex space $V_\lambda \otimes \mathbb{C}^n$.

Proof. Consider again the branching rule for representations of $U(\mathfrak{gl}(n))$

$$V_\lambda \otimes \mathbb{C}^n \simeq \bigoplus_{\lambda' : \lambda \nearrow \lambda'} V_{\lambda'} \,.$$

As was mentioned before, the action of the algebra $U(\mathfrak{gl}(n-1)) \subset U(\mathfrak{gl}(n))$ on both sides of the above expression is also invariant. It follows from (2.15) that both sides split into direct sums of irreducible $U(\mathfrak{gl}(n-1))$-modules and, what's more, these modules and their multiplicities must be the same on both sides. Thus, thinking of the LHS of the above expression as of a $U(\mathfrak{gl}(n-1))$-module, we have, using (2.15)

$$\begin{aligned} V_\lambda \otimes \mathbb{C}^n &\simeq \big(\bigoplus_{\mu : \mu \preceq \lambda} V_\lambda^\mu \big) \otimes \left(\mathbb{C}^{n-1} \oplus \mathbb{C} \right) \\ &\simeq \bigoplus_{\mu : \mu \preceq \lambda} \left[\left(V_\lambda^\mu \otimes \mathbb{C}^{n-1} \right) \oplus \left(V_\lambda^\mu \otimes \mathbb{C} \right) \right] \\ &\simeq \bigoplus_{\mu : \mu \preceq \lambda} \left[\left(V_\lambda^\mu \otimes \mathbb{C}^{n-1} \right) \oplus V_\lambda^\mu \right] . \end{aligned} \tag{5.34}$$

On the other hand, the RHS under the action of $U(\mathfrak{gl}(n-1))$ decomposes as follows

$$\bigoplus_{\lambda' : \lambda \nearrow \lambda'} V_{\lambda'} \simeq \bigoplus_{(\lambda' : \lambda \nearrow \lambda')} \bigoplus_{(\mu' : \mu' \preceq \lambda')} V_{\lambda'}^{\mu'} \,. \tag{5.35}$$

For any $\hat{\mu}$ such that $\mu \nearrow \hat{\mu}$ let

$$\bigoplus_{\lambda'} V_{\lambda'}^{\hat{\mu}} := V(\hat{\mu}) \,,$$

where the sum is over all the modules $V_{\lambda'}^{\hat{\mu}}$ appearing on the right-hand side of (5.35) with partitions λ' such that $\lambda \nearrow \lambda'$ and $\hat{\mu} \preceq \lambda'$; $V(\hat{\mu})$ is just a direct sum of all the copies of $V_{\hat{\mu}}$ when $V_\lambda \otimes \mathbb{C}^n$ is considered as a representation of $U(\mathfrak{gl}(n-1))$. Now consider the branching rule of the module $V_\lambda^\mu \otimes \mathbb{C}^{n-1}$ appearing in (5.34)

$$V_\lambda^\mu \otimes \mathbb{C}^{n-1} \simeq V_\mu \otimes \mathbb{C}^{n-1} \simeq \bigoplus_{\mu'' : \mu \nearrow \mu''} V_{\mu''} .$$

Modules appearing in the decomposition above are all mutually orthogonal and in particular each $V_{\mu''} \subset V_\lambda^\mu \otimes \mathbb{C}^{n-1}$ with $\mu'' \neq \hat{\mu}$ is orthogonal to $V(\hat{\mu})$, since $V(\hat{\mu})$ is just a direct sum of several copies of irreducible modules $V_{\lambda'}^{\hat{\mu}} \simeq V_{\hat{\mu}}$. Thus, using (5.29)

$$\begin{aligned}
\sum_{\lambda'} tr(P_{V_{\lambda'}^{\mu'}} P_{V_\lambda^\mu \otimes \mathbb{C}^{n-1}}) &= tr(P_{V(\mu')} P_{V_\lambda^\mu \otimes \mathbb{C}^{n-1}}) \\
&= dim\big(V(\mu') \cap (V_\lambda^\mu \otimes \mathbb{C}^{n-1})\big) \\
&= dim(V_{\mu'}) = |\mathcal{K}_{\mu'}| ,
\end{aligned}$$

where the last equality follows from the fact that there is one and only one copy of $V_{\mu'}$ in $V_\lambda^\mu \otimes \mathbb{C}^{n-1}$ (in fact, as was mentioned in Ch. 2, each $V_{\mu'} \subset V_\lambda^\mu \otimes \mathbb{C}^{n-1}$, for all μ''s such that $\mu \nearrow \mu'$, has multiplicity one), and all the copies of $V_{\mu'}$ in (5.34) are contained in $V(\mu')$. □

Proof of Proposition 5.13. For all $(\lambda, \mu), (\lambda', \mu') \in \mathcal{W}_{n,n-1}$ with $\mu \neq \mu'$ and *any* admissible choice of ξ we have

$$\begin{aligned}
Q^{(n,2)}((\lambda,\mu),\mu') &= \sum_{\lambda'} Q^{(n,2)}((\lambda,\mu),(\lambda',\mu')) \\
&= \sum_{\lambda'} \sum_{\xi'} P^{(n)}((\lambda,\mu,\xi),(\lambda',\mu',\xi')) \qquad \text{by Prop. 5.12 eq. (5.32)} \\
&= \frac{1}{n} \sum_{\lambda'} \sum_{\xi'} \sum_{k=2}^{n} |\langle e_{(\mu,\xi)}^{\lambda} \otimes e_k, e_{(\mu',\xi')}^{\lambda'} \rangle|^2 ,
\end{aligned}$$

where the summation over e_k starts from $k = 2$ since $|\langle e_{(\mu,\xi)}^{\lambda} \otimes e_1, e_{(\mu',\xi')}^{\lambda'} \rangle|^2$ corresponds to the probability of a particle jumping only in the first row (and hence having $\mu = \mu'$, which we assumed not to be the case). Noting that ξ is arbitrary, the above is equal to

$$
\begin{aligned}
&\frac{1}{n}\frac{1}{|\mathcal{K}_\mu|}\sum_{\lambda'}\sum_{\xi'}\sum_{\xi}\sum_{k=2}^{n}|\langle e^{\lambda}_{(\mu,\xi)}\otimes e_k, e^{\lambda'}_{(\mu',\xi')}\rangle|^2\\
=&\frac{1}{n}\frac{1}{|\mathcal{K}_\mu|}\sum_{\lambda'} tr(P_{V^{\mu'}_{\lambda'}}P_{V^{\mu}_{\lambda}\otimes\mathbb{C}^{n-1}}) \qquad \text{by Lemma 5.14}\\
=&\frac{1}{n}\frac{|\mathcal{K}_{\mu'}|}{|\mathcal{K}_\mu|}=A^{(n-1)}(\mu,\mu')\ .
\end{aligned}
$$

Finally one calculates $A^{(n-1)}(\mu,\mu)=1-\sum_{j=1}^{n-1}A^{(n-1)}(\mu,\mu+e_j)=1/n$. $\square$

We see that the marginal process $\mathcal{Y}$ is behaving as a version of $(n-1)$-dimensional Dyson's random walk with the possibility of no jump at any time $k\geq 0$.

This concludes our analysis of the Markov chain $\mathcal{Z}$ in the GC-cone. In the following section we will discover an interesting link between the process $(\mathcal{X},\mathcal{Y})$ formed by the top two rows of $\mathcal{Z}$ and the Hermitian minor process.

5.4 An informal argument showing convergence of generators

Let $(\mathcal{X}(k),\mathcal{Y}(k);k\geq 0)=(\mathcal{X}_i(k),\mathcal{Y}_j(k);1\leq i\leq n,1\leq j\leq n-1,k\geq 0)$ be the $\mathcal{W}_{n,n-1}$-valued Markov chain with transition probabilities given by

$$
Q^{(n,2)}((x,y),(x+e_i,y))=\frac{1}{n}w^{(1)}\left(\begin{array}{c|c} x+e_i & i\\ y & \end{array}\right)^2,
$$

$$
Q^{(n,2)}((x,y),(x+e_i,y+e_j))=\frac{1}{n}\frac{|\mathcal{K}_{y+e_j}|}{|\mathcal{K}_y|}w^{(2)}\left(\begin{array}{c|c} x+e_i & i\\ y+e_j & j\end{array}\right)^2 .
$$

This is the process we considered at the end of the previous section. We would like to find the connection between the discrete interlaced Markov chain $(\mathcal{X},\mathcal{Y})$ and the Markov process with generator (5.9) formed by the top two rows of the Hermitian minor process considered in Section 5.1.3. One notices that since, by construction precisely one $\mathcal{X}$-particle jumps at any time step, we are losing one degree of freedom (that is, essentially, one dimension in our $(2n-1)$-dimensional process). This forces the particles in the top row of $(\mathscr{X},\mathscr{Y})$ to be correlated in such a way that renders it impossible to obtain an n-point Dyson Brownian motion, which we are hoping for, as a diffusion limit of the marginal process $\mathscr{X}$ (think about the $n=2$ situation, in which case $\mathscr{X}_1(k)$ and $\mathscr{X}_2(k)$ have covariance -1 at any time $k \geq 1$). This prompts us to consider the *Poissonised* version of the process. Let $\mathscr{V} = (\mathscr{X}(t),\mathscr{Y}(t); t \geq 0) = (\mathscr{X}_i(t),\mathscr{Y}_j(t); 1 \leq i \leq n, 1 \leq j \leq n-1, t \geq 0)$ be a continuous-time Markov chain with values in $\mathscr{W}_{n,n-1}$ and the following state-dependent transition intensities

$$\mathscr{Q}((x,y),(x+e_i,y)) := \mathscr{Q}_i(x,y) = w^{(1)} \left(\begin{array}{c|c} x+e_i & i \\ y & \end{array} \right)^2 ,$$

$$\mathscr{Q}((x,y),(x+e_i,y+e_j)) := \mathscr{Q}_{ij}(x,y) = \frac{|\mathscr{K}_{y+e_j}|}{|\mathscr{K}_y|} w^{(2)} \left(\begin{array}{c|c} x+e_i & i \\ y+e_j & j \end{array} \right)^2 .$$

The process $\mathscr{V}$ has the generator $\mathscr{L}f(x,y) = \sum_{i,j}[f(x+e_i,y+e_j)-f(x,y)]\mathscr{Q}_{ij}(x,y) + \sum_i [f(x+e_i,y)-f(x,y)]\mathscr{Q}_i(x,y)$.

By Proposition 5.9 and a calculation in the proof of Proposition 5.13, it follows that the marginal processes $(\mathscr{X}(t); t \geq 0)$ and $(\mathscr{Y}(t); t \geq 0)$, are Poisson counting processes with values in $\mathscr{W}_n$, resp. $\mathscr{W}_{n-1}$, and state-dependent intensities

$$\sum_{j=1}^{n-1} \mathscr{Q}_{ij}(x,y) + \mathscr{Q}_i(x,y) = \frac{|\mathscr{K}_{x+e_i}|}{|\mathscr{K}_x|} := \mathscr{Q}_i^X(x), \quad 1 \leq i \leq n , \tag{5.36}$$

resp.

$$\sum_{i=1}^{n} \mathscr{Q}_{ij}(x,y) = \frac{|\mathscr{K}_{y+e_j}|}{|\mathscr{K}_y|} := \mathscr{Q}_j^Y(y), \quad 1 \leq j \leq n-1 \ . \tag{5.37}$$

No two particles on the same layer (i.e. any two elements of $\mathscr{X}$ or $\mathscr{Y}$) jump simultaneously. However, elements from different layers can. Moreover, a $\mathscr{Y}$-particle must always jump in conjunction with an $\mathscr{X}$-particle, whilst an $\mathscr{X}$-particle can jump on its own.

We define the rescaled process $\mathscr{V}^{(m)} = (\mathscr{X}^{(m)},\mathscr{Y}^{(m)})$ with values in $\mathbb{W}_{n,n-1} =$

$\{(x,y) \in \mathbb{R}^n \times \mathbb{R}^{n-1} : y \preceq x\}$ by letting

$$\mathscr{X}_i^{(m)}(t) = \frac{\mathscr{X}_i(mt) - mt}{\sqrt{m}} \quad \text{and} \quad \mathscr{Y}_j^{(m)}(t) = \frac{\mathscr{Y}_j(mt) - mt}{\sqrt{m}} ,$$

for $1 \le i \le n$, $1 \le j \le n-1$ and all $t \ge 0$. Denote by $\mathbb{P}_{v_0}^m$ be the law of $\mathscr{V}^{(m)}$ started at $v_0 \in \frac{1}{\sqrt{m}}\mathbb{W}_{n,n-1}$.

The object matter of the present section is the follow-up to Conjecture 5.3

Conjecture 5.15. *Let $\mathbb{P}_{v_0}$ denote the unique solution to the martingale problem associated to the generator* (5.9) *and started at $v_0 = o$ or $v_0 \in \mathbb{W}'_{n,n-1}$, where o denotes the origin and $\mathbb{W}'_{n,n-1} = \{(x,y) \in \mathbb{R}^n \times \mathbb{R}^{n-1} : y \prec x\}$ is the interior of $\mathbb{W}_{n,n-1}$. Then*

$$\lim_{m\to\infty} \mathbb{P}_{v_0^m}^m = \mathbb{P}_{v_0} ,$$

where $v_0^m \in \frac{1}{\sqrt{m}}\mathbb{W}'_{n,n-1}$ for all $m \ge 1$, and $\lim_{m\to\infty} v_0^m = v_0$.

The above conjecture states that we expect an appropriately scaled Poissonised version of the process $(\mathscr{X},\mathscr{Y})$ of the top two rows of the Markov chain $\mathscr{Z}$ discussed in previous section to converge in law to the joint process of the top two rows of the Hermitian minor process, thus identifying a discrete version of the latter. We make a first step towards proving the conjecture by showing that the generator of $\mathscr{V}^{(m)}$ converges to $\mathscr{G}$ as m tends to infinity. We argue informally as follows.

For any $m \in \mathbb{N}$ $\mathscr{V}^{(m)}$ is the process with the generator

$$\begin{aligned}
\mathscr{L}^m f(x,y) = &\sum_i [f(x + \frac{1}{\sqrt{m}}e_i, y) - f(x,y)] m\mathscr{Q}_i(\sqrt{m}x, \sqrt{m}y) \\
&+ \sum_{i,j} [f(x + \frac{1}{\sqrt{m}}e_i, y + \frac{1}{\sqrt{m}}e_j) - f(x,y)] m\mathscr{Q}_{ij}(\sqrt{m}x, \sqrt{m}y) \\
&- \sum_i \partial_{x_i} f(x,y)\sqrt{m} - \sum_j \partial_{y_j} f(x,y)\sqrt{m} ,
\end{aligned}$$

where f is any bounded twice-continuously differentiable function with support on the interior of $\mathbb{W}_{n,n-1}$, and, as before, we write $\partial_{x_i} f$ for $\frac{\partial}{\partial x_i} f$, $\partial^2_{x_i y_j} f$ for $\frac{\partial^2}{\partial x_i \partial y_j} f$ and so on.

Now, a multidimensional version of the Taylor approximation series states that for large m we have

$$\frac{1}{2}\partial^2_{x_i x_i} f(x,y) \approx (f(x+\frac{1}{\sqrt{m}}e_i, y) - f(x,y))m - \partial_{x_i} f(x,y)\sqrt{m}\ ,$$

and

$$\begin{aligned}&\frac{1}{2}\partial^2_{x_i x_i} f(x,y) + \partial^2_{x_i y_j} f(x,y) + \frac{1}{2}\partial^2_{y_j y_j} f(x,y)\\ &\qquad \approx (f(x+\frac{1}{\sqrt{m}}e_i, y+\frac{1}{\sqrt{m}}e_j) - f(x,y))m - \partial_{x_i} f(x,y)\sqrt{m} - \partial_{y_j} f(x,y)\sqrt{m}\ .\end{aligned}$$

We now rewrite the expression for $\mathscr{L}^m f(x,y)$ as follows

$$\begin{aligned}\mathscr{L}^m f(x,y) &= \sum_i \left[[f(x+\frac{1}{\sqrt{m}}e_i, y) - f(x,y)]m - \partial_{x_i} f(x,y)\sqrt{m} \right] \mathscr{Q}_i(\sqrt{m}x, \sqrt{m}y)\\ &\quad + \sum_{i,j} \Big[[f(x+\frac{1}{\sqrt{m}}e_i, y+\frac{1}{\sqrt{m}}e_j) - f(x,y)]m\\ &\quad -\partial_{x_i} f(x,y)\sqrt{m} - \partial_{y_j} f(x,y)\sqrt{m} \Big] \mathscr{Q}_{ij}(\sqrt{m}x, \sqrt{m}y)\\ &\quad + \sum_i \partial_{x_i} f(x,y)\sqrt{m} \Big[\sum_j \mathscr{Q}_{ij}(\sqrt{m}x, \sqrt{m}y) + \mathscr{Q}_i(\sqrt{m}x, \sqrt{m}y) - 1 \Big]\\ &\quad + \sum_j \partial_{y_j} f(x,y)\sqrt{m} \Big[\sum_i \mathscr{Q}_{ij}(\sqrt{m}x, \sqrt{m}y) - 1 \Big]\ .\end{aligned}$$

For large enough m, using the Taylor approximation above and relations (5.36) and (5.37), we obtain

$$\begin{aligned}\mathscr{L}^m f(x,y) &\approx \sum_i \frac{1}{2}\partial^2_{x_i x_i} f(x,y)\mathscr{Q}^X_i(\sqrt{m}x) + \sum_j \frac{1}{2}\partial^2_{y_j y_j} f(x,y)\mathscr{Q}^Y_j(\sqrt{m}y)\\ &\quad + \sum_{i,j} \partial^2_{x_i, y_j} f(x,y)\mathscr{Q}_{ij}(\sqrt{m}x, \sqrt{m}y)\\ &\quad + \sum_i \partial_{x_i} f(x,y)\sqrt{m}\left(\mathscr{Q}^X_i(\sqrt{m}x) - 1\right) + \sum_j \partial_{y_j} f(x,y)\sqrt{m}\left(\mathscr{Q}^Y_j(\sqrt{m}y) - 1\right)\ .\end{aligned}$$

Establishing the limit above as m tends to infinity requires us studying the convergence of the reduced Wigner coefficients, essentially constituting the transition intensities for $\mathscr{V}^{(m)}$.

Lemma 5.16.

$$\lim_{m\to\infty} \mathcal{Q}_i^X(\sqrt{m}x) = 1, \quad \lim_{m\to\infty} \mathcal{Q}_i^Y(\sqrt{m}y) = 1\ , \tag{5.38a}$$

$$\lim_{m\to\infty} \mathcal{Q}_{ij}(\sqrt{m}x, \sqrt{m}y) = \prod_{k\neq j} \frac{x_i - y_k}{y_j - y_k} \prod_{k\neq i} \frac{x_k - y_j}{x_k - x_i}\ , \tag{5.38b}$$

$$\begin{aligned} \lim_{m\to\infty} \sqrt{m}\left(\mathcal{Q}_i^X(\sqrt{m}x) - 1\right) &= \sum_{k\neq i} \frac{1}{x_i - x_k}\ , \\ \lim_{m\to\infty} \sqrt{m}\left(\mathcal{Q}_j^Y(\sqrt{m}y) - 1\right) &= \sum_{k\neq j} \frac{1}{y_j - y_k} \end{aligned} \tag{5.38c}$$

for all $x \in \mathbb{W}_n$*,* $y \in \mathbb{W}_{n-1}$ *and* $(x,y) \in \mathbb{W}'_{n,n-1}$*,* $1 \le i \le n$ *and* $1 \le j \le n-1$*. (Recall that for any* $n \ge 1$ $\mathbb{W}_n = \{x \in \mathbb{R}^n : x_1 > x_2 > \ldots > x_n\}$*.)*

Proof. We start with (5.38a). Using an expression for $\mathcal{K}_x$ in terms of the Vandermonde determinant (see eqn. (2.11)), we write

$$\mathcal{Q}_i^X(\sqrt{m}x) = \frac{|\mathcal{K}_{\sqrt{m}x+e_i}|}{|\mathcal{K}_{\sqrt{m}x}|} = \frac{\triangle(\sqrt{m}x + e_i)}{\triangle(\sqrt{m}x)} = \prod_{k\neq i} \frac{\sqrt{m}(x_i - x_k) - i + k + 1}{\sqrt{m}(x_i - x_k) - i + k} \to 1$$

as $m \to \infty$, for all $x \in \mathbb{W}_n$ and $1 \le i \le n$. In a similar fashion one verifies that $\lim_{m\to\infty} \mathcal{Q}_i^Y(\sqrt{m}y) = 1$, $y \in \mathbb{W}_{n-1}$, $1 \le j \le n-1$.

Now, using the explicit expressions for the reduced Wigner coefficients (2.19), we have for (5.38b)

$$\mathcal{Q}_{ij}(\sqrt{m}x, \sqrt{m}y) = \frac{|\mathcal{K}_{\sqrt{m}y+e_j}|}{|\mathcal{K}_{\sqrt{m}y}|} w^{(2)} \left(\begin{array}{c|c} \sqrt{m}x + e_i & i \\ \sqrt{m}y + e_j & j \end{array} \right)^2$$

$$= \prod_{k\neq j} \frac{\sqrt{m}(y_j - y_k) - j + k + 1}{\sqrt{m}(y_j - y_k) - j + k} \prod_{k\neq i} \frac{\sqrt{m}(x_k - y_j) - k + j}{\sqrt{m}(x_k - x_i) - k + i} \prod_{k\neq j} \frac{\sqrt{m}(x_i - y_k) - i + k + 1}{\sqrt{m}(y_j - y_k) - j + k + 1}$$

$$\to \prod_{k\neq j} \frac{x_i - y_k}{y_j - y_k} \prod_{k\neq i} \frac{x_k - y_j}{x_k - x_i} \quad \text{as } m \to \infty$$

for $1 \le i \le n$, $1 \le j \le n-1$ and $(x,y) \in \mathbb{W}'_{n,n-1}$.

Finally we look at (5.38c):

$$\begin{aligned}
\sqrt{m}\left(\mathscr{Q}_i^X(\sqrt{m}x)-1\right) &= \sqrt{m}\left(\frac{|\mathscr{K}_{\sqrt{m}x+e_i}|}{|\mathscr{K}_{\sqrt{m}x}|}-1\right)\\
&= \sqrt{m}\left(\prod_{k\neq i}\left(1+\frac{1}{\sqrt{m}(x_i-x_k)-i+k}\right)-1\right)\\
&= \sum_{k\neq i}\frac{\sqrt{m}}{\sqrt{m}(x_i-x_k)-i+k}+o(1/\sqrt{m}) \to \sum_{k\neq i}\frac{1}{x_i-x_k} \quad \text{as } m\to\infty
\end{aligned}$$

for all $1\le i\le n$, $x\in\mathbb{W}_n$. One finds the limit for $\sqrt{m}\left(\mathscr{Q}_i^Y(\sqrt{m}y)-1\right)$ for all $y\in\mathbb{W}_{n-1}$ in a similar manner. □

It follows that

$$\begin{aligned}
\lim_{m\to\infty}\mathscr{L}^m f(x,y) = {}&\sum_i \frac{1}{2}\partial^2_{x_ix_i}f(x,y)+\sum_j\frac{1}{2}\partial^2_{y_jy_j}f(x,y)\\
&+\sum_i\sum_{k\neq i}\frac{1}{x_i-x_k}\partial_{x_i}f(x,y)+\sum_j\sum_{k\neq j}\frac{1}{y_j-y_k}\partial_{y_j}f(x,y)\\
&+\sum_{i,j}\prod_{k\neq j}\frac{x_i-y_k}{y_j-y_k}\prod_{k\neq i}\frac{x_k-y_j}{x_k-x_i}\partial^2_{x_iy_j}f(x,y)
\end{aligned}$$

for all $(x,y)\in\mathbb{W}'_{n,n-1}$. We write $\mathscr{L}=\lim_{m\to\infty}\mathscr{L}^m$. Now that we have calculated the limiting generator of $\mathscr{V}^{(m)}$, we need to verify that it is indeed the same as the generator (5.9). Clearly, all is left is to check is that the mixed coefficients of the two differential operators are the same; because of the constants γ_j^2's in (5.9), it isn't immediately obvious that it should be the case.

Lemma 5.17. *Let* $(\widetilde{a}_{x_iy_j}; 1\le i\le n, 1\le j\le n-1)$ *and* $(a_{x_iy_j}; 1\le i\le n, 1\le j\le n-1)$ *be the mixed coefficients of the generator* $\mathscr{G}$ (5.9) *and the generator* $\mathscr{L}$ *we found above, respectively (we have switched from notation* (λ,μ) *to that of* (x,y) *for* $\mathscr{G}$*). Then* $\widetilde{a}_{x_iy_j}(x,y)=a_{x_iy_j}(x,y)$ *for all* $(x,y)\in\mathbb{W}'_{n,n-1}$, $1\le i\le n$ *and* $1\le j\le n-1$.

Proof. Recall that for each $1\le j\le n-1$ $\{v_1,\ldots,v_{n-1}\}$, where $v_{ij}=(\widetilde{a}^{1/2}_{x_1y_j},\ldots,\widetilde{a}^{1/2}_{x_ny_j})$, is an orthonormal basis of a random $(n-1)$-dimensional subspace of $\mathbb{C}^n$ (see proof of Thm. 5.2, p. 157). In particular, the coefficients of $\mathscr{G}$ satisfy $\sum_i\widetilde{a}_{x_iy_j}=1$ for all

$1 \le j \le n-1$. At the other hand notice that by the calculation above $a_{x_i y_j}(x,y) = \lim_{m\to\infty} \mathcal{Q}_{ij}(\sqrt{m}x, \sqrt{m}y)$, for any $(x,y) \in \mathbb{W}'_{n,n-1}$, and so

$$\sum_i a_{x_i y_j}(x,y) = \lim_{m\to\infty} \sum_i \mathcal{Q}_{ij}(\sqrt{m}x, \sqrt{m}y) = \lim_{m\to\infty} \mathcal{Q}^Y_j(\sqrt{m}y) = \lim_{m\to\infty} \frac{|\mathcal{K}_{\sqrt{m}y+e_j}|}{|\mathcal{K}_{\sqrt{y}}|} = 1 \; .$$

Furthermore, one easily verifies that

$$\frac{\widetilde{a}_{x_i y_j}(x,y)}{\widetilde{a}_{x_k y_j}(x,y)} = \prod_{k\neq j} \frac{x_i - y_k}{x_p - y_k} \frac{\prod_{k\neq p}(x_k - x_p)}{\prod_{k\neq i}(x_k - x_i)} \frac{x_p - y_j}{x_i - y_j} = \frac{a_{x_i y_j}(x,y)}{a_{x_k y_j}(x,y)}$$

for all $(x,y) \in \mathbb{W}'_{n,n-1}$, $1 \le i,k \le n$.

Coupling this with the fact that $\sum_i \widetilde{a}_{x_i y_j} = \sum_i a_{x_i y_j} = 1$ for all $1 \le j \le n-1$ we deduce that in fact $\widetilde{a}_{x_i y_j} = a_{x_i y_j}$ for all $1 \le i \le n$, $1 \le j \le n-1$. □

Finally we can conclude the limiting generator $\mathcal{L}$ is indeed equal to $\mathcal{G}$ (5.1).

Appendix

Proof of Lemma 5.5. We prove the lemma by induction. In the case $n = 2$ identity (5.19) is verified by an easy calculation. Suppose now it holds for some $n \in \mathbb{N}$. Then for $x, y \in \mathbb{R}^{n+1}$ and any $m \in \{1, ..., n+1\}$

$$\det\left(\frac{1}{x_i - y_j}\right)_{1\le i,j\le n+1} = (-1)^{\alpha} \sum_{k=1}^{n+1} (-1)^{m+k} \frac{1}{x_k - y_m} \det\left(\frac{1}{x_i - y_j}\right)_{\substack{i\neq k\\ j\neq m}} =$$
$$\frac{1}{\prod_{1\le i,j\le n+1}(x_i - y_j)} \sum_{k=1}^{n+1} (-1)^{m+k+1} \prod_{j\neq m}(x_k - y_j) \prod_{i\neq k}(x_i - y_m) \triangle(x^{(k)}) \triangle(y^{(m)}) \; , \tag{5.39}$$

where $x^{(k)}$ is a vector x with k^{th} element deleted an $y^{(m)}$ is defined in a similar manner. Because the above is true for any $1 \le m \le n+1$, we see that the rational function (5.39) is equal to zero, whenever $y_i = y_j$ for any $i \neq j$. Similarly, if we calculate the determinant by expanding along the rows, we conclude that it vanishes whenever

$x_i = x_j$ for any $i \neq j$. Hence

$$\det\left(\frac{1}{x_i - y_j}\right)_{1\leq i,j\leq n+1} \propto \frac{\triangle(x)\triangle(y)}{\prod_{1\leq i,j\leq n+1}(x_i - y_j)}\,.$$

To determine the sign to go in front of the RHS of the expression above, first notice that it is equal to the coefficient of the term $x_1^n y_1^n$, which is the leading term of $\triangle(x)\triangle(y)$. But it follows from (5.39) that the coefficient of $x_1^n y_1^n$ is $(-1)^\alpha(-1)^n$. If n is odd, this coefficient is equal to $(-1)^{\frac{n+1}{2}}$, and if n is even – to $(-1)^{\frac{n}{2}}$. $\square$

Bibliography

[1] M. Adler, E. Nordenstam, and P. van Moerbeke. Consecutive Minors for Dyson's Brownian Motions. *ArXiv e-prints*, July 2010.

[2] J. Baik, P. Deift, and K. Johansson. On the distribution of the length of the longest increasing subsequence of random permutations. *J. Amer. Math. Soc.*, 12(4): 1119–1178, 1999. ISSN 0894-0347. doi: 10.1090/S0894-0347-99-00307-0.

[3] J. Baik, P. Deift, and K. Johansson. On the distribution of the length of the second row of a Young diagram under Plancherel measure. *Geom. Funct. Anal.*, 10(4): 702–731, 2000. ISSN 1016-443X. doi: 10.1007/PL00001635.

[4] M. T. Barlow and P. Protter. On convergence of semimartingales. In *Séminaire de Probabilités, XXIV, 1988/89*, volume 1426 of *Lecture Notes in Math.*, pages 188–193. Springer, Berlin, 1990.

[5] A. O. Barut and R. Raczka. *Theory of group representations and applications*. PWN—Polish Scientific Publishers, Warsaw, 1977.

[6] Yu. Baryshnikov. GUEs and queues. *Probab. Theory Related Fields*, 119(2):256–274, 2001. ISSN 0178-8051. doi: 10.1007/PL00008760.

[7] R. F. Bass. *Diffusions and elliptic operators*. Probability and its Applications (New York). Springer-Verlag, New York, 1998. ISBN 0-387-98315-5.

[8] A. G. Bhatt and R. L. Karandikar. Invariant measures and evolution equations for Markov processes characterized via martingale problems. *Ann. Probab.*, 21(4): 2246–2268, 1993. ISSN 0091-1798.

[9] P. Biane. Marches de Bernoulli quantiques. In *Séminaire de Probabilités, XXIV, 1988/89*, volume 1426 of *Lecture Notes in Math.*, pages 329–344. Springer, Berlin, 1990.

[10] P. Biane. Intertwining of Markov semi-groups, some examples. In *Séminaire de Probabilités, XXIX*, volume 1613 of *Lecture Notes in Math.*, pages 30–36. Springer, Berlin, 1995.

[11] P. Biane. Le théorème de Pitman, le groupe quantique $\mathrm{SU}_q(2)$, et une question de P. A. Meyer. In *In memoriam Paul-André Meyer: Séminaire de Probabilités XXXIX*, volume 1874 of *Lecture Notes in Math.*, pages 61–75. Springer, Berlin, 2006.

[12] P. Biane, P. Bougerol, and N. O'Connell. Littelmann paths and Brownian paths. *Duke Math. J.*, 130(1):127–167, 2005. ISSN 0012-7094. doi: 10.1215/S0012-7094-05-13014-9.

[13] Patrick Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999. ISBN 0-471-19745-9. doi: 10.1002/9780470316962. A Wiley-Interscience Publication.

[14] P. Bougerol and T. Jeulin. Paths in Weyl chambers and random matrices. *Probab. Theory Related Fields*, 124(4):517–543, 2002. ISSN 0178-8051. doi: 10.1007/s004400200221.

[15] N. Bouleau. Semi-martingales à valeurs $\mathbf{R}^d$ et fonctions convexes. *C. R. Acad. Sci. Paris Sér. I Math.*, 292(1):87–90, 1981. ISSN 0151-0509.

[16] N. Bouleau. Formules de changement de variables. *Ann. Inst. H. Poincaré Probab. Statist.*, 20(2):133–145, 1984. ISSN 0246-0203.

[17] A. Bryn-Jones and R. A. Doney. A functional limit theorem for random walk conditioned to stay non-negative. *J. London Math. Soc. (2)*, 74(1):244–258, 2006. ISSN 0024-6107. doi: 10.1112/S0024610706022964.

[18] C. J. Burke and M. Rosenblatt. A Markovian function of a Markov chain. *Ann. Math. Statist.*, 29:1112–1122, 1958. ISSN 0003-4851.

[19] E. Carlen and P. Protter. On semimartingale decompositions of convex functions of semimartingales. *Illinois J. Math.*, 36(3):420–427, 1992. ISSN 0019-2082.

[20] P. Carmona, F. Petit, and M. Yor. Beta-gamma random variables and intertwining relations between certain Markov processes. *Rev. Mat. Iberoamericana*, 14(2): 311–367, 1998. ISSN 0213-2230.

[21] V. Chari and A. Pressley. *A guide to quantum groups*. Cambridge University Press, Cambridge, 1995. ISBN 0-521-55884-0. Corrected reprint of the 1994 original.

[22] J. F. Cornwell. *Group theory in physics. Vol. II*, volume 7 of *Techniques of Physics*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], London, 1984. ISBN 0-12-189802-4.

[23] M. Cranston, W. S. Kendall, and P. March. The radial part of Brownian motion. II. Its life and times on the cut locus. *Probab. Theory Related Fields*, 96(3):353–368, 1993. ISSN 0178-8051. doi: 10.1007/BF01292677.

[24] E. Date, M. Jimbo, and T. Miwa. Representations of $U_q(\mathfrak{gl}(n, \mathbf{C}))$ at $q = 0$ and the Robinson-Shensted [Schensted] correspondence. In *Physics and mathematics of strings*, pages 185–211. World Sci. Publ., Teaneck, NJ, 1990.

[25] M. Defosseux. Orbit measures, random matrix theory and interlaced determinantal processes. *Ann. Inst. Henri Poincaré Probab. Stat.*, 46(1):209–249, 2010. ISSN 0246-0203. doi: 10.1214/09-AIHP314.

[26] J. Dubédat. Reflected planar Brownian motions, intertwining relations and crossing probabilities. *Ann. Inst. H. Poincaré Probab. Statist.*, 40(5):539–552, 2004. ISSN 0246-0203.

[27] E. B. Dynkin. *Theory of Markov processes*. Dover Publications Inc., Mineola, NY, 2006. ISBN 0-486-45305-7. Translated from the Russian by D. E. Brown and edited by T. Köváry, Reprint of the 1961 English translation.

[28] F. J. Dyson. A Brownian-motion model for the eigenvalues of a random matrix. *J. Mathematical Phys.*, 3:1191–1198, 1962. ISSN 0022-2488.

[29] S. N. Ethier and T. G. Kurtz. *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986. ISBN 0-471-08186-8. Characterization and convergence.

[30] W. Feller. Non-Markovian processes with the semigroup property. *Ann. Math. Statist.*, 30:1252–1253, 1959. ISSN 0003-4851.

[31] William Feller. *An introduction to probability theory and its applications. Vol. I.* Third edition. John Wiley & Sons Inc., New York, 1968.

[32] H. Föllmer and P. Protter. On Itô's formula for multidimensional Brownian motion. *Probab. Theory Related Fields*, 116(1):1–20, 2000. ISSN 0178-8051.

[33] H. Föllmer, P. Protter, and A. N. Shiryayev. Quadratic covariation and an extension of Itô's formula. *Bernoulli*, 1(1-2):149–169, 1995. ISSN 1350-7265.

[34] M. Fuhrman and G. Tessitore. Generalized directional gradients, backward stochastic differential equations and mild solutions of semilinear parabolic equations. *Appl. Math. Optim.*, 51(3):279–332, 2005. ISSN 0095-4616. doi: 10.1007/s00245-004-0814-x.

[35] W. Fulton and J. Harris. *Representation theory*, volume 129 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1991. ISBN 0-387-97527-6; 0-387-97495-4. A first course, Readings in Mathematics.

[36] J. R. Giles. *Convex analysis with application in the differentiation of convex functions*, volume 58 of *Research Notes in Mathematics*. Pitman (Advanced Publishing Program), Boston, Mass., 1982. ISBN 0-273-08537-9.

[37] J. Gravner, C. A. Tracy, and H. Widom. Limit theorems for height fluctuations in a class of discrete space and time growth models. *J. Statist. Phys.*, 102(5-6): 1085–1132, 2001. ISSN 0022-4715. doi: 10.1023/A:1004879725949.

[38] N. F Grinberg. Semimartingale decomposition of convex functions of continuous semimartingales by Brownian perturbation. *ArXiv e-prints*, July 2009.

[39] J. Hachigian. Collapsed Markov chains and the Chapman-Kolmogorov equation. *Ann. Math. Statist.*, 34:233–237, 1963. ISSN 0003-4851.

[40] J. Hachigian and M. Rosenblatt. Functions of reversible Markov processes that are Markovian. *J. Math. Mech.*, 11:951–960, 1962.

[41] B. M. Hambly, J. B. Martin, and N. O'Connell. Concentration results for a Brownian directed percolation problem. *Stochastic Process. Appl.*, 102(2):207–220, 2002. ISSN 0304-4149. doi: 10.1016/S0304-4149(02)00177-1.

[42] N. Ikeda and S. Watanabe. *Stochastic differential equations and diffusion processes*, volume 24 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam, 1981. ISBN 0-444-86172-6.

[43] K. Johansson. Shape fluctuations and random matrices. *Comm. Math. Phys.*, 209 (2):437–476, 2000. ISSN 0010-3616. doi: 10.1007/s002200050027.

[44] K. Johansson. Discrete orthogonal polynomial ensembles and the Plancherel measure. *Ann. of Math. (2)*, 153(1):259–296, 2001. ISSN 0003-486X. doi: 10.2307/2661375.

[45] K. Johansson. Discrete orthogonal polynomial ensembles and the Plancherel measure. *Ann. of Math. (2)*, 153(1):259–296, 2001. ISSN 0003-486X. doi: 10.2307/2661375.

[46] K. Johansson. The arctic circle boundary and the Airy process. *Ann. Probab.*, 33 (1):1–30, 2005. ISSN 0091-1798. doi: 10.1214/009117904000000937.

[47] K. Johansson and E. Nordenstam. Eigenvalues of GUE minors. *Electron. J. Probab.*, 11:no. 50, 1342–1371, 2006. ISSN 1083-6489.

[48] Ch. Kassel. Quantum groups. In *Algebra and operator theory (Tashkent, 1997)*, pages 213–236. Kluwer Acad. Publ., Dordrecht, 1998.

[49] M. Katori and H. Tanemura. Nonintersecting Paths, Noncolliding Diffusion Processes and Representation Theory. *ArXiv Mathematics e-prints*, January 2005.

[50] F. P. Kelly. Markovian functions of a Markov chain. *Sankhyā Ser. A*, 44(3):372–379, 1982. ISSN 0581-572X.

[51] W. S. Kendall. The radial part of Brownian motion on a manifold: a semimartingale property. *Ann. Probab.*, 15(4):1491–1500, 1987. ISSN 0091-1798.

[52] W. König. Orthogonal polynomial ensembles in probability theory. *Probab. Surv.*, 2:385–447 (electronic), 2005. ISSN 1549-5787. doi: 10.1214/154957805100000177.

[53] W. König and N. O'Connell. Eigenvalues of the Laguerre process as non-colliding squared Bessel processes. *Electron. Comm. Probab.*, 6:107–114 (electronic), 2001. ISSN 1083-589X.

[54] W. König, N. O'Connell, and S. Roch. Non-colliding random walks, tandem queues, and discrete orthogonal polynomial ensembles. *Electron. J. Probab.*, 7: no. 5, 24 pp. (electronic), 2002. ISSN 1083-6489.

[55] S. Lang. *Analysis II*. Addison-Wesley, 1969.

[56] M. L. Mehta. *Random matrices*, volume 142 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, third edition, 2004. ISBN 0-12-088409-7.

[57] P.-A. Meyer. Un cours sur les intégrales stochastiques. In *Séminaire de Probabilités, X (Seconde partie: Théorie des intégrales stochastiques, Univ. Strasbourg, Strasbourg, année universitaire 1974/1975)*, pages 245–400. Lecture Notes in Math., Vol. 511. Springer, Berlin, 1976.

[58] N. O'Connell. Conditioned random walks and the RSK correspondence. *J. Phys. A*, 36(12):3049–3066, 2003. ISSN 0305-4470. Random matrix theory.

[59] N. O'Connell. A path-transformation for random walks and the Robinson-Schensted correspondence. *Trans. Amer. Math. Soc.*, 355(9):3669–3697 (electronic), 2003. ISSN 0002-9947. doi: 10.1090/S0002-9947-03-03226-4.

[60] N. O'Connell and M. Yor. A representation for non-colliding random walks. *Electron. Comm. Probab.*, 7:1–12 (electronic), 2002. ISSN 1083-589X.

[61] K. R. Parthasarathy. *An introduction to quantum stochastic calculus*, volume 85 of *Monographs in Mathematics*. Birkhäuser Verlag, Basel, 1992. ISBN 3-7643-2697-2.

[62] J. W. Pitman. One-dimensional Brownian motion and the three-dimensional Bessel process. *Advances in Appl. Probability*, 7(3):511–526, 1975. ISSN 0001-8678.

[63] M. Prähofer and H. Spohn. Statistical self-similarity of one-dimensional growth processes. *Phys. A*, 279(1-4):342–352, 2000. ISSN 0378-4371. doi: 10.1016/S0378-4371(99)00517-8. Statistical mechanics: from rigorous results to applications.

[64] P. Protter. *Stochastic integration and differential equations*, volume 21 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 1990. ISBN 3-540-50996-8. A new approach.

[65] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999. ISBN 3-540-64325-7.

[66] R. T. Rockafellar. *Convex analysis*. Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J., 1970.

[67] L. C. G. Rogers. Characterising one-dimensional diffusions using stochastic calculus. *Bull. London Math. Soc.*, 19(2):183–185, 1987. ISSN 0024-6093.

[68] L. C. G. Rogers and J. W. Pitman. Markov functions. *Ann. Probab.*, 9(4):573–582, 1981. ISSN 0091-1798.

[69] L. C. G. Rogers and D. Williams. *Diffusions, Markov processes, and martingales. Vol. 2*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2000. ISBN 0-521-77593-0. Itô calculus, Reprint of the second (1994) edition.

[70] M. Rosenblatt. Functions of Markov processes. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 5:232–243, 1966.

[71] F. Russo and P. Vallois. The generalized covariation process and Itô formula. *Stochastic Process. Appl.*, 59(1):81–104, 1995. ISSN 0304-4149. doi: 10.1016/0304-4149(95)93237-A.

[72] F. Russo and P. Vallois. Itô formula for C^1-functions of semimartingales. *Probab. Theory Related Fields*, 104(1):27–41, 1996. ISSN 0178-8051. doi: 10.1007/BF01303801.

[73] B. E. Sagan. *The symmetric group*, volume 203 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 2001. ISBN 0-387-95067-2. Representations, combinatorial algorithms, and symmetric functions.

[74] B. Simon. *Representations of finite and compact groups*, volume 10 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1996. ISBN 0-8218-0453-7.

[75] R. P. Stanley. *Enumerative combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999. ISBN 0-521-56069-1; 0-521-78987-7. With a foreword by Gian-Carlo Rota and appendix 1 by Sergey Fomin.

[76] D. W. Stroock and S. R. S. Varadhan. *Multidimensional diffusion processes*, volume 233 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1979. ISBN 3-540-90353-4.

[77] C. A. Tracy and H. Widom. Level-spacing distributions and the Airy kernel. *Comm. Math. Phys.*, 159(1):151–174, 1994. ISSN 0010-3616.

[78] W. von Waldenfels. The Markov process of total spins. In *Séminaire de Probabilités, XXIV, 1988/89*, volume 1426 of *Lecture Notes in Math.*, pages 357–361. Springer, Berlin, 1990.

[79] J. Warren. Dyson's Brownian motions, intertwining and interlacing. *Electron. J. Probab.*, 12:no. 19, 573–590 (electronic), 2007. ISSN 1083-6489.

[80] J. Warren and P. Windridge. Some examples of dynamics for Gelfand-Tsetlin patterns. *Electron. J. Probab.*, 14:no. 59, 1745–1769, 2009. ISSN 1083-6489.

[81] D. Williams. *Weighing the odds*. Cambridge University Press, Cambridge, 2001. ISBN 0-521-80356-X; 0-521-00618-X. A course in probability and statistics.